TRANSACTIONS OF
THE
OPTICAL SOCIETY

VOL. XXXI. 1929-30 No. 3.

PRESIDENTIAL ADDRESS

OPTICS IN RADIO TRANSMISSION AND OTHER FRESH FIELDS

BY FRANK TWYMAN, F.R.S., F.INST.P.

13th February, 1930

THERE can be few whose vocation is optical who have not felt at times the narrowness of their sphere of activity. Whether manufacturers, teachers, experimenters or designers we are all liable to feel that we are working within narrow bounds.

We have allowed ourselves to be cribbed, cabined and confined by such definitions as the one which says:

> Optics is a science which treats of light and the phenomena of vision.

It is true that both the art and science of optics were first of all concerned only with that small range of radiation which alone affects the eye: but when in 1800 W. Herschel showed that the power of the visible spectrum to heat bodies was shared by an invisible extension of the spectrum beyond the red end, and when, the next year, Ritter showed that silver chloride was blackened not only by visible light, but by an invisible extension of the spectrum beyond the violet end, the restrictions imposed by the dictionary were removed once for all.

One might spend an evening profitably in considering the optical developments still to be achieved by further exploitation of the visible, infra-red, and ultra-violet radiations which since 1801 have all been regarded as equally optical. For instance, to quote from a recent article by Dr G. E. Hale:

> Large burning lenses and concave mirrors were long ago used to perform chemical experiments and to vaporise metals, and recently Straubel of the Carl Zeiss Company has found it possible to obtain a temperature of nearly 5000° C. (9000° F.) at the focus of a searchlight mirror. As this will instantly vaporise tungsten and is within about 1000° of the surface temperature of the sun, a properly designed "solar" furnace should be of great service in extending the range of our electric furnaces, so successfully used by King and others for the spectroscopic interpretation of astrophysical phenomena.

But it is my purpose to-night to dwell on additions to physical knowledge which have recently been made outside this range, and to show that they, too, are proper subjects for the curricula of teachers of optics, and for the products of optical manufacturers.

Let us define optics afresh as the science that deals with radiation; with the production, control and utilisation of radiation. If the wave theory becomes involved with corpuscles then let us say that corpuscular radiation belongs to us also. After all, optical radiation was held by Newton to be corpuscular, and Hamilton in his famous optical papers developed in connection with optics mathematical expressions which originated in corpuscular ideas. It is true that during the development of his work the wave theory came into vogue, but the form of his expressions remained equally appropriate to the later ideas, and with the indifference to the nature of things peculiar to the mathematician he seems to have been quite undisturbed by the change.

I propose to-night, then, to mention a few territories on the confines of the Optical Empire which can be ours if we do not let them slip from our hands.

The first that comes to mind embraces those radiations which are used for broadcasting, namely, optical radiations which have a wave-length of from 3×10^6 to 3×10^{14} angstroms—I state the wave-lengths in the optical unit (see table below).

Range of wave-lengths covered by different kinds of radiation.

Kind of radiation	A	cm.	km.
Cosmic rays	from ·0002 to ·06	from 2×10^{-12} to 6×10^{-10}	
γ-rays	from ·05 to ·20	from 5×10^{-10} to 2×10^{-9}	
X-rays	from ·10 to 100	from 10^{-9} to 10^{-6}	
Ultra-violet	from 100 to 3900	from 10^{-6} to $3{\cdot}9 \times 10^{-5}$	
Visible	from 3900 to 7600	from $3{\cdot}9 \times 10^{-5}$ to $7{\cdot}6 \times 10^{-5}$	
Infra-red	from 7600 to 420×10^4	from $7{\cdot}6 \times 10^{-5}$ to ·042	
"Wireless" rays	from 3×10^6 to 3×10^{14}	from ·03 to 3×10^8	up to 30

The "Angstrom," A, is a ten-millionth of a millimetre.

These travel with the same speed as our other optical radiations and can by suitable optical means be directed in beams which can show the usual phenomena of diffraction, reflection and interference.

Prof. Appleton* draws attention to the similarity of wireless and ordinary optical phenomena. He notes first that since the aerial is normally a vertical wire the vibrating electricity in it generates waves which are polarised. The periodic electric force in the waves—which corresponds to what we call the vibration direction in optics—is therefore vertical for horizontal transmission along the ground, and is accompanied by a periodic magnetic force at right angles both to the electric force and to the direction of wave propagation. The magnetic force is therefore horizontal and is in phase with the electric force.

If we have two emitting stations of the same wave-length, we have conditions which will give us interference phenomena, as in optics. Some years ago, the

* "Some Large Scale Optical Experiments," E. V. Appleton, F.R.S., *School Science Review*, No. 39, March, 1929. Also *Proc. Roy. Soc.* A, **109** 1925 (621); **113** 1926 (450).

engineers of the B.B.C. tried to work two stations on the same wave-length, and interference maxima and minima were experienced in the region round the stations, the signal variations being specially marked at points equidistant from the two stations. It was found difficult to keep the phase difference between the two stations constant, so that there was a gradual shifting of the interference pattern with time, causing variations of signal intensity.

With only one emitting station, we may experience interference phenomena comparable with those of a Lloyd's mirror system, due to the fact that the ionised region of the upper atmosphere acts like a mirror to wireless waves and sends them back to the ground. At a receiving station situated at about a hundred miles from an ordinary broadcasting station the two sets of waves, one of which travels direct along the ground and the other of which is returned from the atmosphere, are of about equal intensity, and interference phenomena are specially marked. This interference system has been used to give an experimental proof of the existence of this ionised layer and also an estimate of its height.

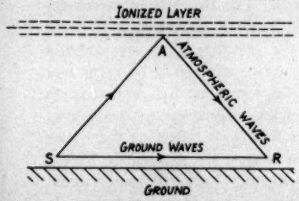


Fig. 1. Interference of direct and reflected wireless waves

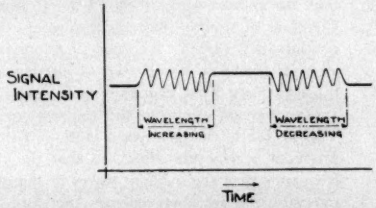


Fig. 2. Interference of direct and reflected wireless waves

Figs. 1 and 2. Illustrating Prof. E. V. Appleton's article on "Some Large Scale Optical Experiments." (*School Science Review*, No. 39, March, 1929.)

Let us consider the simplest possible case of interference between ground waves and atmospheric waves. We may assume as a first approximation that the latter are sharply reflected at the surface of the ionised layer, or Heaviside layer, as it is usually called.

If there is a whole number of wave-lengths path difference between the ground and atmospheric ray tracks, there will be a maximum of intensity at R, while if the path difference is an odd number of half wave-lengths, a minimum will be experienced. Let us suppose that the wave-length of the emitted radiation is slowly and continuously altered. This will produce a succession of interference maxima and minima at R, and if the number n of either is counted, and the initial and final wave-lengths (λ_1 and λ_2 respectively) of the change are known, the difference D between the ground and atmospheric paths may be found. Evidently, if D does not alter with the wave-length, then

$$n = \frac{D}{\lambda_1} - \frac{D}{\lambda_2}.$$

D having been found thus, and the distance from S to R being known, the height of the layer may be found by simple triangulation.

Fig. 2 illustrates a number of maxima produced by such a change of wave-length caused by a continuous frequency change taken at Peterborough on a transmission from the National Physical Laboratory on the 24th May, 1929.

Mean wave-length 400 metres
Mean frequency 750 kc./s.

Wave-length change 10 metres
Frequency change 18·7 kc./s.

Two changes of frequency are of course involved, from λ_1 to λ_2 and back again to λ_1.

There is, however, another way of regarding the maxima and minima. Since the atmospheric path is longer than the ground path, the atmospheric waves received at any instant will have set out from the emitting station a little earlier than the ground waves received at the same instant. The two sets of waves are therefore of slightly different frequency, since the emitted wave-length is being continuously altered. We may therefore regard the signal maxima and minima as beats produced by two sets of electromagnetic waves of slightly different frequency. Very often, especially in the early morning, somewhat less regular beats are experienced even when the frequency emitted by the sending station is constant. This is due to the downward motion of the layer which on reflecting the waves produces a Döppler effect, so that the reflected waves are of slightly higher frequency than the incident waves or the ground waves. The variations of signal intensity that are observed may therefore quite well be considered as beats. It will be remembered that light beats are obtainable with the Michelson Interferometer if we move one of the mirrors parallel to itself. Due to the Döppler effect, the frequency of the beam reflected from the moving mirror is different from that reflected from the fixed mirror, and the moving interference fringes experienced may be regarded as beats.

The wave-length change method of finding the equivalent height of the Heaviside layer has been used in experiments carried out under the auspices of the Radio Research Board of the Department of Scientific and Industrial Research. A special series of transmissions from the National Physical Laboratory has been received at Peterborough and the diurnal variation of transmission studied. During the day-time, when the layer is lower because of solar influence, the down-coming ray is weak, but after sunset recombination of the ions takes place and the under boundary of the layer rises, reaching its maximum height just before dawn. At the same time, the reflection coefficient of the layer is found to increase.

The process by which the waves are deviated by the ionised layer has been much discussed, and although there are still many points that are obscure, it is generally thought that long waves are truly reflected at a discontinuity of ionisation which is sharp compared with a wave-length, while short waves are bent back gradually by a process of refraction

The transmission of the waves emitted by the London B.B.C. Station 2 LO has been the subject of an interesting study by R. H. Barfield, working for the Radio Research Board of the Department of Scientific and Industrial Research. In attempting to account for the attenuation of the waves travelling along the ground, Mr Barfield has examined the possibilities of absorption by buildings, wireless aerials and trees. In the heart of a big city, steel-structured buildings cause marked absorption of the waves, whereas in the suburbs the chief source of attenuation seems to be the large density of wireless aerials tuned to the frequency of the waves. In the country it is found that dense masses of trees, acting as untuned aerials, also reduce the strength of the waves passing over them.

The case of selective absorption by tuned wireless aerials is of particular interest in that it is comparable in some ways to the case of anomalous dispersion in physical optics. The theory given by Planck for the absorption and dispersion of light waves due to the presence, in the medium, of a large number of electrical vibrators with a characteristic frequency, may, with very little change, be adapted to explain the chief features of the effect noticed by Mr Barfield.

The growth of material appliances for producing and using these radiations has been marvellously rapid under the stimulus of popular appetite, but although the electrical wireless companies have had large sums of money at their disposal for the past thirty years, it seems to have taken them about twenty years to arrive at beam transmission. Is it not reasonable to suppose that it would have been one of the first ideas which we opticians would have applied if we had not been musing in our cloistered seclusion?

When the first account of Marconi's proposals was read before the Institute of Electrical Engineers I distinctly remember Professor Silvanus Thompson pointing out the disadvantages of a system of signalling which spread without control in all directions.

May I ask you to pardon my interpolating an account of how an optician proposed to deal with sound in an attempt to control its direction?

Hearing and seeing are sister senses, in that they are almost universally regarded as the only proper media for the conveyance of high aesthetic pleasure. But more to the point is the fact that there is a considerable analogy between the expressions for the propagation of sound and light.

Foghorns have probably been used in these islands since the time of the Romans. If I remember correctly, there is at Dover Castle a horn supposed to date from Roman times which may have supplemented the beacon there to guide the galleys through the Channel fogs. The anxiety of the mariner, wrapped in woolly fog, and unable to tell from which direction comes the sound of the warning whistle, shout, or bell, must be an experience renewed with but slight emotional variation throughout thousands of years.

It occurred to me early in 1914, very naturally as an optician, that it should be possible to produce a sound beam by providing a source of sound consisting of a row of syrens uniformly disposed in a straight line, and emitting synchronous sound waves.

SOUND BEAMS

British Patent No. 4797/14
U.S.A. Patent No. 1287595

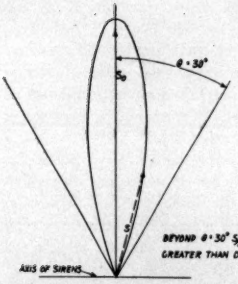


Fig. 3. The curve indicates the directional distribution of sound intensity. No. of sirens = 10.
$b/\lambda = 0{\cdot}2$ (viz. distance between successive sirens is 0·2 times the wave-length of the note employed).
For 224 vibrations per second, wave-length = 5 feet, total length of line of sirens = 9 feet.

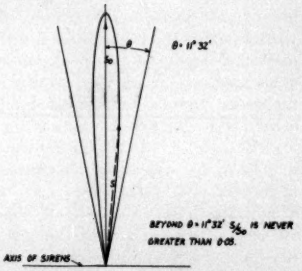


Fig. 4. The curve indicates the directional distribution of sound intensity. No. of sirens = 10.
$b/\lambda = 0{\cdot}5$ (viz. distance between successive sirens is 0·5 times the wave-length of the note employed).
For 224 vibrations per second, wave-length = 5 feet, total length of line of sirens = 22 feet 6 inches.

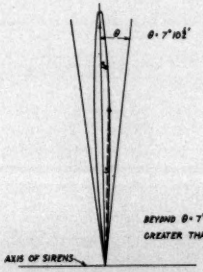


Fig. 5. The curve indicates the directional distribution of sound intensity. No. of sirens = 10.
$b/\lambda = 0{\cdot}8$ (viz. distance between successive sirens is 0·8 times the wave-length of the note employed).
For 224 vibrations per second, wave-length = 5 feet, total length of line of sirens = 36 feet.

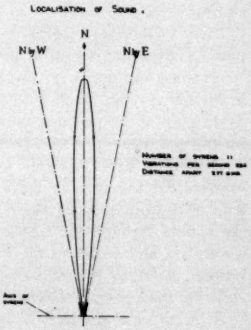


Fig. 6.

The idea was never realised, but Dr L. Silberstein kindly worked out for me the degrees of localisation that might be anticipated, and these are illustrated in Figs. 3, 4, 5 and 6.

Now let us revert to beam wireless transmission, and see the angle from which it is considered by one whose interests have till recently been largely optical. (A description will be found in W. E. Williams' British Patent Application 317301.)

One of the principal objects of a wireless directional antenna is to confine the major portion of the energy of the transmitter into a narrow beam, and the vertical wire radiators of such a transmitter are analogous to the illuminated slits or grooves of a diffraction grating. Just as in the diffraction grating, the more radiators (with a given spacing between each) the narrower is the resulting beam.

In the ordinary diffraction grating, the spacing is greater than a wave-length and so we have two or three or more principal maxima, or, as we call them, orders of spectra. This would be wasteful and uneconomical in wireless so the order must be restricted to a single one, and hitherto this has been accomplished by keeping the radiators a fraction of a wave-length apart (usually about a quarter). This entails a large, costly, and difficult arrangement, since each radiator has to be separately spaced, tuned and phased.

Now it is well known that in Michelson's Echelon Grating in "single order" position we have only *one* principal maximum, although the lateral separation between the radiators in the optical case may be several thousand wave-lengths. One asks oneself what is the essential principle in Michelson's arrangement that makes this possible. The answer is that here all the directions of principal maxima caused by the collection of widely separated elements (except one direction) coincide with directions of zero amplitude for each individual element, so that there is nothing to add up for these directions.

Suppose your optical radiator A (Fig. 7) has a principal maximum in direction θ_1 and a minimum (zero) in directions θ_2, θ_3, and so on. All one has to do then, is to add other radiators such as B, C, etc., in such a way that the lateral separation, X, in a direction perpendicular to θ_1 is

$$X = \frac{\lambda}{\sin \alpha}$$

and all the orders except that in the θ_1 direction will then vanish, for under this condition the first order spectrum will then be in direction θ_2 (making an angle α with direction θ_1) where the intensity is a minimum.

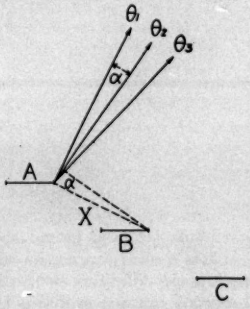


Fig. 7. Echelon Radiation.

As an example of how this kind of reasoning may be applied to beam transmission we may consider a commonly used French system of antennae (Fig. 8). The dots denote vertical radiators with a separation of about a quarter

wave-length. All in a horizontal row are equiphased and the phases of successive rows of antennae are chosen so that all waves reinforce in the direction of the arrow.

Now consider a single column A, B. This by itself will confine the radiation within a fairly narrow beam, and it is easy from ordinary diffraction grating theory to calculate the angle α. (For a row 32 wave-lengths long, it works out that α is of the order of 15° so that sin α = approximately $\frac{1}{4}$.)

Then if the next row CD, instead of being about $\frac{1}{4}\lambda$ from AB, is put, as in Fig. 9, at a distance

$$X = 4\lambda$$

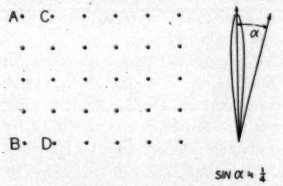


Fig. 8. French system of "wireless" antennae for beam transmission.

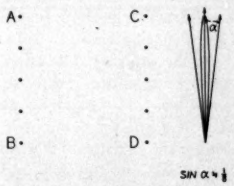


Fig. 9. Williams' system of "wireless" antennae for beam transmission.

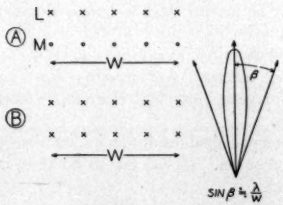


Fig. 10. Arrangements for cancelling backward radiation in beam transmission.

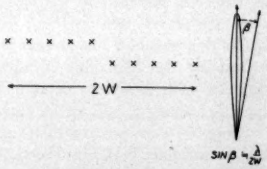


Fig. 11. Williams' system of cancelling backward radiation in beam transmission.

we have the same type of radiation pattern as a single row, but the radiation is confined within about half the former angle, since with $X = 4\lambda$ A and C will neutralise each other, when $X \sin \alpha = \lambda/2$, viz. when $\sin \alpha = 1/8$. Therefore the intensity becomes four times as great. It is sheer waste to pack columns with $\frac{1}{4}\lambda$ spacing between them.

To prevent backward radiation, two methods have hitherto been used.

In one ×, ×, × [see Fig. 10 (A)] are equiphased energised tuned wires, and o, o, o, a reflector system $\frac{1}{4}\lambda$ behind, also tuned but not energised.

This method depends on an induction effect which does not give exactly the right phase difference, so that if the separation be chosen nearly $\frac{1}{4}\lambda$ so that the backward radiation cancels out the forward, the two do not fully reinforce each other. On the other hand, if the separation be chosen so that the forward radiation

is fully enhanced, the backward radiation is not perfectly annulled. Any small change of wave-length doubles the error, since in the forward direction the reflected wave from L to M and back to L has to cover the distance twice. The width of the beam is from our ordinary diffraction grating theory approximately

$$\sin \beta \fallingdotseq \frac{\lambda}{W}.$$

The other method [Fig. 10 (B)] is simply to energise two sets at 90° phase and $\frac{1}{4}\lambda$ separation, the half angle β being the same as before. This method is wasteful of energy, so that the previous system is more popular.

Now we know from optics that the displacement of a *parallel* beam has no effect provided we are not concerned with the energy distribution in the *immediate vicinity* of the transmitter, but, for instance, on the other side of the earth. We can therefore replace the above arrangement Fig. 10 (B) by that shown in Fig. 11. Here we are keeping to the Echelon principle. Our beam will now be twice as narrow, while retaining all the good points of the previous arrangements.

This has the advantages over the reflector system shown in Fig. 10 (A) in that the tuned circuits are not required; when the wave-length alters, by slightly varying the phasing we can partially compensate for the fact that the separation is no longer $\frac{1}{4}\lambda$. Further, at the correct wave-length phase difference and separation, the two forward waves *exactly* reinforce and the two back radiations *exactly* cancel.

Having now dealt with a large scale phenomenon—the emission of wireless radiations from a system of antennae—let us consider the small scale one represented by the diffraction of X-rays from crystals.

It was in 1912 that von Laue showed that X-rays could be diffracted by a crystal, and thus deduced that they are of the same nature as light (Fig. 12).

Fig. 12 shows the Laue photograph obtained by passing a narrow beam of X-rays through a crystal of nickel sulphate.

Following von Laue's work Sir William (then Prof. W. H.) Bragg and W. L. Bragg* showed that the Laue photographs could be conveniently interpreted as due to the reflection of the X-rays in such planes within the crystal as are rich in atoms. This led them to use the cleavage planes of crystals as mirrors, and their researches soon showed them how to apply these ideas to determine not only the wave-lengths of X-radiations, but in many instances the exact distribution of atoms in a crystal.

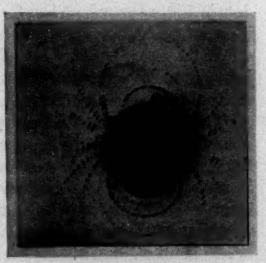


Fig. 12. Laue Photograph, Nickel Sulphate. (*X-rays and Crystal Structure*, Sir W. H. and W. L. Bragg, 1924.)

The early work of the Braggs was soon followed by that of Moseley† who with appropriately modified apparatus, recorded photographically the distribution of X-radiations in

* *Proc. Roy. Soc.* A **88** 1913 (428).

† *Phil. Mag.* **26** 1913 (1024).

accordance with their wave-lengths, thus producing a true spectrogram of the distinctive X-rays emitted by various elements.

X-ray spectrography has since then followed both courses of development with important results, investigations directed like Moseley's on the wave-lengths of the X-rays emitted by the various elements, and those whose purpose has been to determine how atoms are arranged in crystals. Both lines of investigation have afforded brilliant results, the first in elucidating the nature of the radiating atom, the second in placing on a sure foundation our knowledge of the structure of molecules.

I have always for convenience (though without any authority for doing so) called this second line of work "X-ray crystallography."

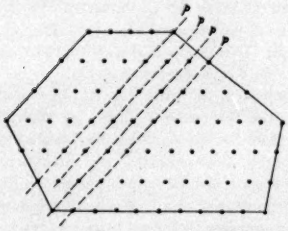


Fig. 13. From *X-rays and Crystal Structure*, by Sir W. H. and W. L. Bragg, p. 9, 1924. (London: Messrs Bell & Sons.)

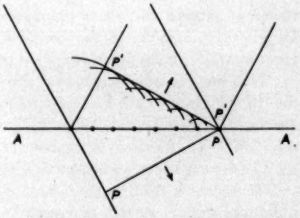


Fig. 14. From *X-rays and Crystal Structure*, by Sir W. H. and W. L. Bragg, p. 8, 1924. (London: Messrs Bell & Sons.)

The crystal structure revealed by X-rays may be illustrated by Fig. 13. The dots in Fig. 13 may be taken for purposes of illustration as representing the atoms in a crystal, the distance between neighbouring atoms being in the case of rock salt of the order 3×10^{-8} cm. Observe the similarity to the system of dots shown just now which represented a system of radio antennae. It is not just a superficial resemblance; there is a close connection between the physical realities represented.

There is a distinction in that the antennae are long, whereas the atoms are, at any rate as far as present considerations are concerned, points. But this distinction is lost when we are at the distance from the antennae which are of interest in radio transmission. Further, the antennae are sources of radiation, while the atoms are acting as secondary sources, at any rate in the diffraction of X-rays which we are now considering, each atom scattering the radiations which it receives more or less in all directions.

Let us suppose that we have a series of particles, Fig. 14, which all lie in one plane *AA*, these particles representing the atoms which scatter the waves. When a pulse passes over these atoms, each emits a diffracted pulse which spreads spherically around it.

Fig. 14 illustrates the result of the passage of the pulse *PP* over the atoms in

the plane AA. The arcs of circles represent the pulses sent out by atoms in the plane. It is clear that all the diffracted wavelets touch a reflected wave $P'P'$: in fact we have only repeated Huygens' construction for the wave front reflected from a plane surface. It does not matter how the particles are arranged on the plane AA, as long as they lie exactly on that plane.

Thus, when the pulse passes over a set of particles which lie in a plane, the diffracted pulses combine to form a wave front which obeys the usual laws of reflection from the plane.

Reverting now to Fig. 13 it is evident that its atoms may be considered as arranged in planes in an infinite number of ways.

It is the natural arrangement of the particles in planes which gives rise to the plane faces of crystals, just as the form of the crystal in Fig. 13 is a polygon bounded by straight lines. It has been seen that if a wave passes over the crystal all the particles in any plane combine to reflect it with respect to that plane. Therefore, if we choose any one way of arranging the crystal particles in planes, corresponding to the lines p, p, p, p, in Fig. 13, and then find the direction in which a wave would be reflected by these parallel planes, this direction will be one in which to expect an interference maximum. There is an infinite number of ways of choosing such planes, but the most effective planes are those which are most thickly studded with particles, and which have correspondingly a wide spacing between successive planes. Study of crystal forms has shown that the natural faces which crystals exhibit can be explained as being on the whole those of a simple type, on which the particles are distributed with the greatest density. Therefore one might say, to sum up, that when a pulse passes over the crystal its scattered energy is concentrated into definite beams, and that these beams may be regarded as faint reflections of the pulses on planes parallel to possible faces in the interior of the crystal.

The reflection does not depend upon the existence of any polished surface on the outside of the crystal, it depends upon an arrangement of planes within. It is this difference between what is here called the "reflection" and the true reflection of ordinary light on surfaces which makes the analogy at first somewhat puzzling. When we talk of the X-rays being reflected by the crystal, it must be borne in mind that the term is only used in order to simplify the conception of the effect. There can be no such thing as a surface reflection of X-rays. The surface reflection of light is an effect concerned with the external form of the reflecting body alone. As long as the reflecting film is more than a few wave-lengths thick the full reflection is obtained. But when X-rays fall on a crystal face the first few layers of atoms cannot diffract any appreciable proportion of the rays, for experiment shows that the rays must pass through some millions of layers before the X-ray beam is appreciably absorbed. Reflection takes place on all the layers, the only limitation being that if a layer is too deep within the body of the crystal, the absorption by the superincumbent layers reduces its contribution to an amount that is negligible. To the comparatively long waves of light the atomic structure is so fine grained that the medium is practically continuous. The X-rays are so short, on the other hand, that a crystal is to them a series of widely separated and

regularly arranged particles, each of which diffracts a very small proportion of their energy.

So far we have considered the reflection of a single pulse. We may now proceed to consider the reflection of a regular train of waves. Each plane reflects the wave train as a wave train, but when the reflected trains are in the same phase, that is to say, are so arranged that they fit on to each other exactly, the reflected energy is far greater than if the condition is imperfectly fulfilled, even if the want of fit is exceedingly small.

It will be readily seen that for X-radiation of any particular wave-length, this will occur at a particular angle which corresponds with the distance between the planes p, p, p, p, containing a large number of atoms, such as those shown in Fig. 15. In short, when the crystal is turned into a certain orientation depending on the wave-length of the X-rays, we shall get a powerful reflection of X-rays, and in all other directions a very feeble reflection. By confining the impulse to a very narrow beam and by slowly rotating the crystal we therefore get a true spectrum. The

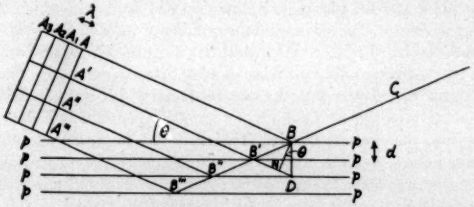


Fig. 15. Bragg's condition for strong reflection of X-rays from crystals.

word "spectrum" in this connection means exactly what it usually does, namely, an arrangement of the radiations from a source in order of their wave-lengths. However, it will be observed that an X-ray spectrum can give two kinds of information. First, like the ordinary optical spectrum, it can serve as a method of analysing by means of its emission spectrum any substance which, by cathode ray bombardment, is caused to emit X-rays. Secondly, it can perform the remarkable function of enabling the distribution of atoms within crystals to be determined. Anyone unfamiliar with this subject will find it most clearly explained in Sir W. H. and Prof. W. L. Bragg's *X-rays and Crystal Structure* from which I have taken this description.

Figs. 16 and 17 show examples of the first kind of work. The first is taken from Moseley's paper previously cited, while the second shows the definition that is obtainable with a modern instrument of the type designed by Prof. Laby.

Figs. 18 and 19 illustrate the determination by X-ray crystallography of the distribution of atoms within a crystal. Fig. 18 shows the positions in which the radiations of a given known wave-length are distributed by reflection at the most

important planes of the crystal, those planes, namely, which contain large numbers of atoms.

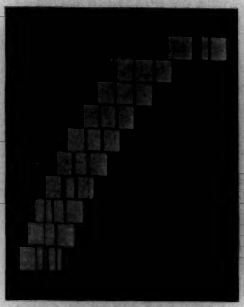


Fig. 16. X-ray Spectra. (H. G. J. Moseley. *Phil. Mag.* **26** 1913 (1025).)

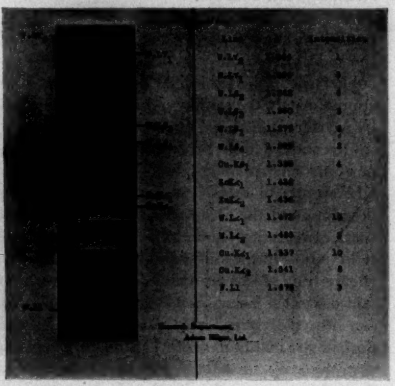


Fig. 17.

Fig. 18. X-ray Crystallograms of KCl and NaCl. (A. W. Hull.)

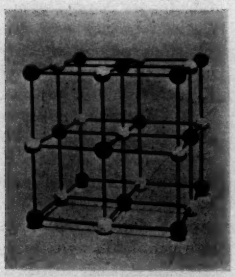


Fig. 19. Sir William and Prof. W. L. Bragg's crystal models. (1) Rocksalt type.

I had hoped to be able to show you a figure of a crystal model, and at the same time a reproduction of an actual X-ray crystallogram which had been instrumental in the elucidation of the same. The nearest I have been able to get, however, are the two crystallograms obtained by Hull for potassium and sodium chlorides, Fig. 18, and in the next slide, Fig. 19, the model of the rocksalt type of crystal worked out by Sir W. H. Bragg.

Although this crystal structure was obtained by an ionisation X-ray spectrograph, yet the two crystallograms in Fig. 18 would yield sufficient data for constructing such a model.

A simple cube with an atom at each corner would explain the photograph of potassium chloride, while the different intensities of certain lines in the case of the sodium chloride would necessitate the stipulation of the larger face-centred cubic structure. If the light sodium atoms were replaced by potassium atoms of approximately the same weight as the chloride atoms the simpler structure would be recovered.

Recently a diffraction grating spectrograph has been designed* which will not only take photographs of lines in the far ultra-violet region, but will also take spectra of soft X-rays without readjustment. It will be seen how, in the description of the ordinary X-ray spectrograph which I have just given, when we speak of reflection from crystals we must understand that it is not ordinary specular reflection from the surface which is in question, but rather a sum of the diffracted energy from atoms regularly distributed within the crystal in such a way that the said diffracted radiation adds up to a utilisable amount. Soft X-rays, however, can be regularly reflected, and in the Thibaud X-ray spectrograph, which will deal in an identical manner with all radiations from ultra-violet ones of 3000 angstroms up to soft X-rays of 10 angstroms. Since beyond 1850 angstroms these rays are entirely absorbed by air, the instrument works of course *in vacuo*, and the sources of radiation—a spark for the ultra-violet, and a hot cathode X-ray tube for the X-ray region—must also be evacuated, and form an integral part of the instrument when it is in use.

The conditions under which the X-rays are reflected are peculiar and very interesting. It was shown by Compton that the refractive index of glass for the X-rays is slightly less than unity. It follows, therefore, that at grazing incident X-rays may be totally reflected. Dr Thibaud in his spectrograph uses a grating at nearly grazing incidence, and in this way has been able to get spectrograms of the soft X-ray region with exposures not unreasonably long.

More recently similar work has been done by Siegbahn using a speculum metal grating.

In connection with this, it is desirable that I should explain briefly the manner in which the X-ray crystallographic examination is carried out in the examination of powders, since that is the form in which most substances present themselves in practice, and not in the form of single crystals.

In a fine powder of a crystalline material (and most materials are crystalline from this point of view), crystals are present in every variety of orientation. It is not necessary, therefore, to alter the angle of incidence of the X-rays on the surface of a powder which is to be examined, since, at any rate if the grains are fine, there will always be a sufficient number of crystals turned into the right orientation to fulfil Bragg's condition for copious reflection.

Of course, since only a small proportion of the crystals are operative in this

* Thibaud, *Phys. Zeit.* **29** May 1928 (241, 1); *Journ. Amer. Opt. Soc.* **17** August 1928 (145).

reflection, one has the concomitant disadvantage that a much longer exposure must be allowed than is required where one has a single perfect crystal turned into exactly the right orientation.

We will now go over to a type of radiation which is undeniably corpuscular in its nature, the description of which I draw from Aston's *Isotopes*.* In 1886 Goldstein described a type of electrical discharge in vacuum tubes, which has since become known as a positive ray because it consists of streams of positively charged atoms or molecules.

In a vacuum tube discharge a stream of electrons at a very high velocity is sent off by the cathode. The electrons being extremely light rapidly attain a high velocity. Such discharges of electrons are called cathode rays, and when they strike a suitable target that target becomes a source of X-rays. If, however, in their progress through the tube they collide with an atom of the residual gases, it may cause the detachment of an electron from that atom. The remaining part of the atom, therefore, becomes positively charged, and under the action of the electric field moves towards the cathode from which the electrons are streaming.

If the field is strong, and the pressure so low that the ionised atom does not collide with the other atoms too frequently, it will ultimately attain a high speed, and the surface of the cathode will therefore be under a continuous hail of positively charged particles varying from the lightest atom to the heaviest molecule present in the discharge tube. They will be of all velocities up to a maximum which depends on the total potential applied to the electrodes and on the mass of the atom or molecule.

If a hole be pierced in the cathode, the positive rays pass through the hole and form a stream of atoms and molecules varying both in mass and velocity, and to this stream the name of "Positive Rays" is given.

The positively charged atoms or molecules are immensely more massive than the electron, for the mass of the lightest atom, that of hydrogen, is about 1845 times that of the electron. Thus the positive rays do not acquire the same speed as the cathode rays.

It is originally to Sir J. J. Thomson that we owe the method of sorting out the positive rays and obtaining from them a great variety of information. He passed the rays through a very narrow tube, thus producing a very fine beam, and sorted them out by electric and magnetic fields. The construction of one of the types of apparatus used is indicated in Fig. 20. The discharge by which the rays are made takes place in a large flask *A* similar to an ordinary X-ray bulb of about 1½ litres capacity. The aluminium cathode *B* is placed in the neck of the bulb, and has a hole in it, continued as an extremely fine-bore tube, through which the positive rays pass. The anode is an aluminium rod *D*, which is generally placed for convenience in a side tube. In order to ensure a supply of the gas under examination, a steady stream is allowed to leak in through an exceedingly fine glass capillary tube *E*, and after circulating through the apparatus is pumped off at *F*.

A voltage of from 30,000 to 50,000 volts is used, and, if the air pressure in *A*

* Aston, *Isotopes* (London: Edward Arnold & Co.).

is kept right, positive rays are produced which, under the influence of this high potential, attain high speeds as they fly towards the cathode. Thus positive rays falling axially pass right through the fine tube and emerge as a narrow beam.

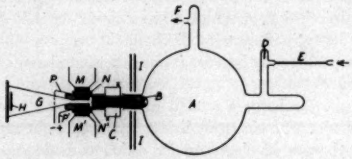


Fig. 20. Positive ray apparatus. [From Aston's *Isotopes*, Fig. 3, p. 26 (London: Edward Arnold & Co.).]

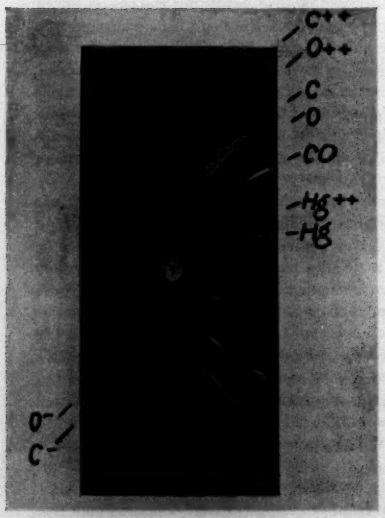


Fig. 21. Photograph of typical positive ray parabolas. [From Aston's *Isotopes*, Fig. 2, Plate 1 (London: Edward Arnold & Co.).]

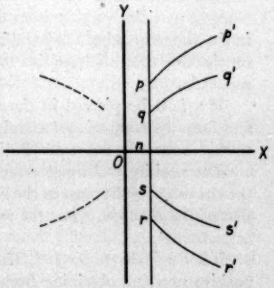


Fig. 22. Positive ray parabolas. [From Aston's *Isotopes*, Fig. 4, p. 28 (London: Edward Arnold & Co.).]

The arrangement for sorting out the rays consists of pieces of soft iron PP', which are placed between the poles MM' of a powerful electro-magnet. Although

P and P' constitute the pole pieces of the magnet, they are electrically insulated from it by thin sheets of mica NN', so that they can be raised to any desired potential difference. Thus the rays in their passage are submitted both to an electrical and to a magnetic field. If there is no field between the plates PP', the beam of rays will strike the screen of the highly exhausted camera G at a point in line with the fine tube, which we may call the undeflected spot. If an electrical field is now applied between the plates a particle will be deflected in the plane of the paper. In the same way, if the electrical field is removed and a magnetic field applied between P and P', the particle will be deflected at right angles to the plane of the paper. Since particles of any one kind have every variety of velocity up to a maximum, it is easy to see that they will be variously deflected, and as a matter of fact they will be so deflected as all to lie on a parabola. If, for instance, all the particles consist of singly ionised atoms of mercury, they will all lie on the parabola marked Hg shown in Fig. 21. If, during ionisation, two electrons are split off instead of one, the resulting positive ray will have a double charge, but calculation shows

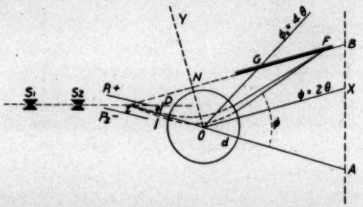


Fig. 23. Diagram of mass spectrograph. [From Aston's *Isotopes*, Fig. 8, p. 45 (London: Edward Arnold & Co.).]

that its behaviour will remain quite indistinguishable from that of a normal ray of half its mass. Parabolas due to doubly ionised Hg, O, and C, will be seen on the lantern slide. Confining our attention, however, to particles (whether atoms or molecules) having a single charge, it is found that for every mass there is a distinct parabola.

The method of comparing masses will be seen from Fig. 22. Suppose pp' and qq' are two such parabolas as appeared on the last slide, and let a line drawn parallel to the magnetic axis OY meet these parabolas in p and q respectively and the electric axis OX in n. The ratio of the masses is in inverse proportion to the squares of the intercepts pn and qn; that is to say, the masses of the particles can be compared by merely measuring lengths on the photographs.

We are now ready to consider a further development of this instrument known as the Aston Mass Spectrograph. Aston discovered that it was possible to produce a focussing effect on any one kind of charged particles, and this purpose is achieved by the mass spectrograph illustrated in Fig. 23.

The rays after arriving at the cathode face pass through two very narrow parallel slits $S\,1$, $S\,2$, and the resulting thin ribbon is spread out into an electric spectrum

by means of the parallel plates P_1, P_2. They are thus deflected by amounts varying according to their mass and velocity, but it is found that in the first order of approximation these rays may all be taken as radiating from a source Z, half way through the field on the line S_1, S_2. A group of these rays selected by means of the diaphragm D, is allowed to pass between the parallel poles of a magnet, these poles for simplicity being illustrated on the diagram by a circle. If the field between these poles is uniform, and of such a size as to bend the rays in the opposite direction to the electrical field, it is found that, with a sufficient degree of approximation, the rays of any one kind will all lie on a straight line ZF drawn through Z. Each

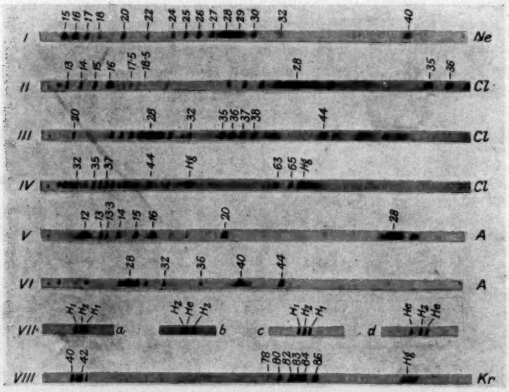


Fig. 24. Mass Spectra. [From Aston's *Isotopes*, Plate III (London: Edward Arnold & Co.).]

kind of particle is distributed along this line ZF to a position which indicates its mass, in the beautiful way shown in Fig. 24, and it is from these spectra that Aston confirmed and extended Soddy's discoveries of the isotopes—that is, varieties of atoms which, while chemically identical, differ in atomic weight.

You have listened very patiently while I have strung together these casually selected examples of a wider optics.

There can be little new in them to most of you, except that they have been brought together and—may I say—formally proclaimed as optical by your President. If you endorse that annexation, and if you have acquired with me a sense of a freer scope available for us opticians, I need not excuse either my subject or its handling.

The passages of which my address chiefly consists referring to the work of Sir W. H. and W. L. Bragg, Aston, Appleton and Williams are given almost entirely in the words of these workers.

IMAGERY AROUND A SKEW RAY

By T. SMITH, M.A., F.Inst.P.
(Optics Department, National Physical Laboratory)

MS. received, 31*st January*, 1929. *Read and discussed*, 13*th March*, 1930.

ABSTRACT. The one-one correspondence between rays in the object and image spaces of an optical instrument only yields a corresponding relation between points, even for a narrow ray pencil, in exceptional circumstances. The usual criterion by which conjugate points are defined fails unless neighbouring rays intersect in both spaces. In general this condition is not satisfied, and the criterion is therefore extended to include the points of nearest approach of non-intersecting rays. It is shown that pairs of conjugate points on a given chief ray determined according to this definition are coincident for a fan of near rays with one degree of freedom, that is the light path between the conjugate points is constant to the second order for the routes corresponding to all rays of this fan, and that the enlarged definition is applicable to all near rays. The relation connecting conjugate points is of the same form as for the simpler cases generally recognised, and may be determined by parallel projection through fixed points depending on the chief ray and on the particular fan under consideration. Apart from the exceptional case of complete anastigmatism due to the coincidence of the projective centres for all the fans, there can at most be two pairs of conjugate points free from astigmatism. Three conditions must be satisfied for stigmatic imagery at a given pair of points.

The pencil around any skew ray has ten degrees of freedom, and the fundamental coefficients which determine the imagery completely are elements of a square matrix of the fourth order. Formulae are given for the construction of these matrices for refraction at each surface and for the transference from one surface to another. The matrix for the complete system is the product of the matrix factors for the successive elementary events taken in their proper order. Each refracting surface may be replaced by the osculating surface of the second degree at the point of refraction of the chief ray. It is assumed that the curvature of the surfaces is everywhere finite and continuous.

A more natural matrix for refraction in space of three dimensions would be of the sixth order. Such matrices are derived from those of the fourth order. For theoretical investigations and perhaps also for routine numerical work, these matrices are preferable to those of lower order since they contain no quantities which are not relevant to the problem. The coefficients of the eikonal and of the characteristic function can be derived from the elements of either type of matrix. Fifteen independent relations between the coefficients of these matrices are obtained in two distinct forms.

Part I. Conjugate Points

The Geometrical Definition of Conjugate Points

According to the usual definition of geometrical optics two points are said to be conjugate to one another when there exist several rays which initially pass through one of the points and, after deviation by reflection or refraction at the surfaces of the instrument under consideration, pass through the other point. The rays in the initial

and final media are supposed to be of unlimited length in both directions, so that both real and virtual points are included.

The definition limits very severely our ideas of conjugate points. If we consider a ray (hereafter called a chief ray) refracted at a single spherical surface, it is well known that a neighbouring ray diverging from a point on the incident portion of the chief ray will not in general intersect the chief ray again after refraction. On account of the symmetry of a sphere the conditions for re-intersection are unusually favourable. Intersections will be much less frequent when the rays have been deviated at a number of surfaces in succession. The limitations imposed by the definition are illustrated by the fact that, though elaborate discussions are frequently given of the configuration of a pencil surrounding a chief ray which lies entirely in a plane containing the axis of an axially symmetrical instrument, the properties of a pencil surrounding a skew chief ray are not usually considered in optical treatises.

Reverting to the single spherical surface, the double intersection takes place only

(1) when both points of intersection lie on the surface of the sphere, or

(2) (in the case of refraction only) the points of intersection fall on the aplantic surfaces, or

(3) the two rays are equidistant from the centre of the sphere, or

(4) the two rays lie in a plane passing through the centre of the sphere.

In a general discussion on conjugate points (1) and (2) are of trivial importance. The second ray in (3) may be brought into coincidence with the chief ray by rotation about the line joining their first intersection to the centre of the sphere. It follows that the join of conjugate points in this case passes through the centre of the sphere. Conjugate points of this type are called secondary conjugate foci, and the centre of the sphere is the secondary projective centre. When pencils of small aperture are considered, the secondary rays in each medium lie in a plane normal to that in which the chief ray is refracted.

When the two rays are coplanar throughout it is obvious that they will intersect twice. It was shown by Thomas Young that the conjugate points of the chief ray thus determined are independent of the particular additional ray chosen, provided its inclination to the chief ray is small. Conjugate points of this type are known as primary conjugate foci, and the plane of the rays is the primary plane. It is normal to the plane containing the chief ray and a secondary ray.

Thomas Young also showed that the join of conjugate primary foci passes through a fixed point, the position of which however depends on the chief ray of the pencil. His method of finding this point is essentially to drop a perpendicular CQ from C, the centre of the sphere, to PP', the straight line connecting P and P', the feet of the perpendiculars from C to the two portions of the chief ray. The primary perspective centre is Q, the foot of this perpendicular.

It is obviously a necessary condition for complete freedom from astigmatism that all perspective centres should coincide. When this condition is not satisfied there can at most be only two positions on any ray where there is no astigmatism, viz. where the two portions of the chief ray meet one another and where they are

met by the straight line CQ. These points are those already considered under (1) and (2) respectively above.

The extension of these results from one to several surfaces depends on the identity at successive refractions of the planes for which conjugate points and perspective centres exist. We may either project through the various centres in succession, and thus by a finite series of operations find the conjugate of any given object point, or we may find (most simply by a direct geometrical construction) two perspective centres for the complete instrument, one for the object space and the other for the image space, such that the line joining the object point to the former is parallel to that joining the latter to the image point.

Generalisation of the Geometrical Definition of Conjugate Points, and the Physical Significance of the Relation

To extend the concept of conjugate points we must adopt a criterion which agrees with the existing definition when successive rays meet but does not break down when no intersection occurs. Now as a point moves along a straight line its distance from any other straight line not parallel to the first will vary. If the point starts from a sufficiently remote neighbourhood the distance will decrease continuously to zero if the lines intersect and to a minimum value if they do not meet; thereafter the distance will increase. The point of one line from which the distance to the other is a minimum may obviously be regarded as a generalisation of the point of intersection of straight lines lying in the same plane. If we apply this generalisation to a chief and a neighbouring ray we shall obviously obtain two definite points on the former, one in the object space and one in the image space. We will define these as the conjugate points of the chief ray determined by that particular neighbouring ray. Clearly the two rays determine conjugate points without ambiguity, no matter how many reflections or refractions occur.

To be of optical importance this extended geometrical definition must correspond to some simple property connected with the waves of light. Let $P\mathfrak{p}$ be the shortest distance between the two rays in the object space, and $P'\mathfrak{p}'$ the corresponding distance in the image space, where P and P' are the conjugate points of the chief ray. Let a plane wave normal to the chief ray start from P at the same time as a plane wave normal to the auxiliary ray starts from $\mathfrak{p}$. As $P\mathfrak{p}$ is normal to both rays, it is the line in which the two wave fronts intersect initially. After passing through the instrument the waves may still be treated as plane in the neighbourhood of P' and $\mathfrak{p}'$ provided we are only considering small transverse distances. With this reservation (which in effect has been assumed throughout) we see that each of the two waves will pass through $\mathfrak{p}'$ at the same moment as it goes through P'; this evidently involves simultaneous passage of both waves through the line $P'\mathfrak{p}'$. We infer that the time taken by light to travel from the one line to the other along certain paths near the chief ray is independent of the terminal points of these paths provided they are near P and P'. In other words there will be agreement in the phase of wavelets reaching such points as P and P' from points on a definite section of the wave front

to which the chief ray is a normal. This is exactly the property which forms the basis of the physical definition of conjugate points. We may therefore conclude that this extension of the concept of conjugate points will prove of value.

Refraction at a Single Surface

From the nature of the new definition it should be as readily applicable to surfaces of any shape as to spherical surfaces. We will first verify this by considering light refracted in the neighbourhood of the origin of the surface

$$2x = Ay^2 + 2Byz + Cz^2 + \text{terms of higher order in } y \text{ and } z,$$

using the criterion of equality of path.

Consider the length $\mathscr{L}$ of the path from the point

$$(\rho\cos\phi,\quad \rho\sin\phi\cos\alpha,\quad \rho\sin\phi\sin\alpha)$$

via (x, y, z) to $\qquad (\rho'\cos\phi', \rho'\sin\phi'\cos\beta, \rho'\sin\phi'\sin\beta)$.

We have
$$\begin{aligned}\mathscr{L} &= \mu'\left[(\rho'\cos\phi' - x)^2 + (\rho'\sin\phi'\cos\beta - y)^2 + (\rho'\sin\phi'\sin\beta - z)^2\right]^{\frac{1}{2}}\\
&\quad - \mu\left[(\rho\cos\phi - x)^2 + (\rho\sin\phi\cos\alpha - y)^2 + (\rho'\sin\phi\sin\alpha - z)^2\right]^{\frac{1}{2}}\\
&= \mu'\rho' - \mu\rho - x(\mu'\cos\phi' - \mu\cos\phi)\\
&\quad - y(\mu'\sin\phi'\cos\beta - \mu\sin\phi\cos\alpha) - z(\mu'\sin\phi'\sin\beta - \mu\sin\phi\sin\alpha)\\
&\quad + \tfrac{1}{2}(x^2 + y^2 + z^2)\left(\frac{\mu'}{\rho'} - \frac{\mu}{\rho}\right)\\
&\quad - \tfrac{1}{2}\left[\frac{\mu'}{\rho'}(x\cos\phi' + y\sin\phi'\cos\beta + z\sin\phi'\sin\beta)^2\right.\\
&\qquad \left. - \frac{\mu}{\rho}(x\cos\phi + y\sin\phi\cos\alpha + z\sin\phi\sin\alpha)^2\right]\\
&\quad + \text{terms of higher order.}\end{aligned}$$

From the equation of the surface x is of smaller magnitude than y or z, and the physical principles of wave propagation give at once that

$$\mu'\sin\phi'\cos\beta - \mu\sin\phi\cos\alpha = \mu'\sin\phi'\sin\beta - \mu\sin\phi\sin\alpha = 0,$$

or
$$\beta = \alpha,\quad \mu'\sin\phi' = \mu\sin\phi,$$

the geometrical law of refraction. Substituting now for x from the equation of the surface and rearranging, we have

$$\begin{aligned}\mathscr{L} &= \mu'\rho' - \mu\rho\\
&\quad - \tfrac{1}{2}\left[(\mu'\cos\phi' - \mu\cos\phi)(Ay^2 + 2Byz + Cz^2) - \left(\frac{\mu'}{\rho'} - \frac{\mu}{\rho}\right)(y^2 + z^2)\right.\\
&\qquad \left. + \left(\frac{\mu'\sin^2\phi'}{\rho'} - \frac{\mu\sin^2\phi}{\rho}\right)(y\cos\alpha + z\sin\alpha)^2\right]\\
&\quad + \text{terms of higher orders in } y \text{ and } z\\
&= \mu'\rho' - \mu\rho\\
&\quad - \tfrac{1}{2}(y^2 + z^2)\left[\tfrac{1}{2}(\mu'\cos\phi' - \mu\cos\phi)\{A + C + (A - C)\cos 2\theta + 2B\sin 2\theta\}\right.\\
&\qquad \left. - \left(\frac{\mu'}{\rho'} - \frac{\mu}{\rho}\right) + \left(\frac{\mu'\sin^2\phi'}{\rho'} - \frac{\mu\sin^2\phi}{\rho}\right)\cos^2(\alpha - \theta)\right]\\
&\quad + \text{higher terms,}\end{aligned}$$

where $\theta = \tan^{-1}\frac{z}{y}$. Thus, whatever values α and θ may have there is a definite relation of the form

$$\frac{k'}{\rho'} - \frac{k}{\rho} = K,$$

where k, k' and K are constants connecting conjugate points on a chief ray, which holds for all paths near the chief ray meeting the refracting surface in an arbitrarily given section through the point of incidence of the chief ray. This relation is of the form which enables conjugate points to be determined by projection through a fixed point. We infer that the conjugate points for passage through an instrument consisting of many surfaces can be determined by the same process as is used for the more restricted sets of points hitherto recognised. The only precaution to be taken is to select projection points which relate to the same path plane.

Since in general k, k' and K are all functions of θ, the variation in the position of the point conjugate to a fixed point as the plane θ is rotated is of a complicated character. This is clearly the cause of the failure of the primary and secondary conjugate points considered in our previous enquiry* to determine the extent of the astigmatism of a pencil surrounding a skew ray. When the surfaces of the system are spherical the problem is greatly simplified by the vanishing of $A - C$ and B. We may observe that the astigmatism is then measured by the curvature

$$(\mu \sin\phi)^2\left(\frac{1}{\mu'\rho'} - \frac{1}{\mu\rho}\right)$$

and only vanishes when the incidence is normal or when the refracting surface is at equal optical distances from the two conjugate points.

The General Theory

The conclusions we have reached may be easily demonstrated with complete generality with the aid of the eikonal. Let origins be taken in the object space and image space at arbitrarily chosen points on the chief ray, and let the axes of x be arranged to coincide with this ray in both spaces. The eikonal, which represents the length of the stationary optical path between planes through the origins having respectively the direction cosines (L, M, N) and (L', M', N'), may be expanded in the form

$$\begin{aligned}\mathscr{E} = {} & e_0 + e_1M + e_2N + e_3M' + e_4N' \\ & + \tfrac{1}{2}e_{11}M^2 + \tfrac{1}{2}e_{22}N^2 + \tfrac{1}{2}e_{33}M'^2 + \tfrac{1}{2}e_{44}N'^2 \\ & + e_{12}MN + e_{34}M'N' + e_{13}MM' + e_{24}NN' + e_{14}MN' + e_{23}M'N \\ & + \text{terms of higher orders.}\end{aligned}$$

The ray with direction cosines (L, M, N) and (L', M', N') meets the planes through the origins normal to the chief ray in the points

$$\left(0, \frac{1}{\mu}\cdot\frac{\partial\mathscr{E}}{\partial M}, \frac{1}{\mu}\cdot\frac{\partial\mathscr{E}}{\partial N}\right) \text{ and } \left(0, -\frac{1}{\mu'}\cdot\frac{\partial\mathscr{E}}{\partial M'}, -\frac{1}{\mu'}\cdot\frac{\partial\mathscr{E}}{\partial N'}\right)$$

* "Note on skew pencils traversing a symmetrical instrument," *Trans. Opt. Soc.* **30** 1928–9 (130).

respectively. As M, N, M', N' are all continuously diminished to zero this ray must continuously approach and eventually coincide with the chief ray. In other words, $\frac{\partial \mathscr{E}}{\partial M}, \frac{\partial \mathscr{E}}{\partial N}, \frac{\partial \mathscr{E}}{\partial M'}, \frac{\partial \mathscr{E}}{\partial N'}$ must all vanish with M, N, M', N', or

$$e_1 = e_2 = e_3 = e_4 = 0$$

identically.

The eikonal referred to the points $(x, 0, 0)$, $(x', 0, 0)$ is

$$\mathscr{E}' = \mathscr{E} + \mu' L' x' - \mu L x,$$

and thus points on the ray $(L, M, N; L', M', N')$ are (x, y, z), (x', y', z'), where

$$y = \frac{1}{\mu} \cdot \frac{\partial \mathscr{E}}{\partial M} + \frac{M}{L} x, \qquad z = \frac{1}{\mu} \cdot \frac{\partial \mathscr{E}}{\partial N} + \frac{N}{L} x,$$

$$y' = -\frac{1}{\mu'} \cdot \frac{\partial \mathscr{E}}{\partial M'} + \frac{M'}{L'} x', \quad z' = -\frac{1}{\mu'} \cdot \frac{\partial \mathscr{E}}{\partial M'} + \frac{N'}{L'} x'.$$

The distances τ and τ' of these points from the chief ray are given by

$$\tau^2 = y^2 + z^2, \quad \tau'^2 = y'^2 + z'^2, \quad \text{or}$$

$$\mu^2 (M^2 + N^2) \tau^2 = \left(M \frac{\partial \mathscr{E}}{\partial N} - N \frac{\partial \mathscr{E}}{\partial M} \right)^2 + \left\{ \mu x \frac{(M^2 + N^2)}{L} + M \frac{\partial \mathscr{E}}{\partial M} + N \frac{\partial \mathscr{E}}{\partial N} \right\}^2,$$

$$\mu'^2 (M'^2 + N'^2) \tau'^2 = \left(M' \frac{\partial \mathscr{E}}{\partial N'} - N' \frac{\partial \mathscr{E}}{\partial M'} \right)^2 + \left\{ \mu' x' \frac{M'^2 + N'^2}{L'} - M' \frac{\partial \mathscr{E}}{\partial M'} - N' \frac{\partial \mathscr{E}}{\partial N'} \right\}^2.$$

The expressions in the last bracket of each equation vanish when τ and τ' have their minimum values, and of course show that $(x, 0, 0)$ and $(x', 0, 0)$ then lie in the planes through (x, y, z) and (x', y', z') respectively normal to the directions (L, M, N) and (L', M', N'). The length of the path from (x, y, z) to (x', y', z'), viz.

$$\mathscr{E} + \mu' (L' x' + M' y' + N' z') - \mu (L x + M y + N z),$$

is equal to $\mathscr{E}'$ when τ and τ' have their minimum values, i.e. it is equal to

$$\mathscr{E} + \mu' x' - \mu x - \frac{L' (1 - L')}{M'^2 + N'^2} \left(M' \frac{\partial \mathscr{E}}{\partial M'} + N' \frac{\partial \mathscr{E}}{\partial N'} \right)^2 - \frac{L (1 - L)}{M^2 + N^2} \left(M \frac{\partial \mathscr{E}}{\partial M} + N \frac{\partial \mathscr{E}}{\partial N} \right)$$

$$= \mathscr{E} + \mu' x' - \mu x - \frac{1}{2} \left(1 - \frac{1 - L'}{1 + L'} \right) \left(M' \frac{\partial \mathscr{E}}{\partial M'} + N' \frac{\partial \mathscr{E}}{\partial N'} \right)$$

$$- \tfrac{1}{2} \left(1 - \frac{1 - L}{1 + L} \right) \left(M \frac{\partial \mathscr{E}}{\partial M} + N \frac{\partial E}{\partial N} \right).$$

Remembering that first order terms are missing from $\mathscr{E}$, we may write

$$\mathscr{E} = \mathscr{E}_0 + \mathscr{E}_2 + \mathscr{E}_3 + \ldots,$$

$$M \frac{\partial \mathscr{E}}{\partial M} + N \frac{\partial \mathscr{E}}{\partial N} + M' \frac{\partial \mathscr{E}}{\partial M'} + N' \frac{\partial \mathscr{E}}{\partial N'} = 2\mathscr{E}_2 + 3\mathscr{E}_3 + \ldots,$$

where the suffix denotes the order of the term. We then obtain

path from (x, y, z) to (x', y', z')

$$= \mathscr{E}_0 + \mathscr{E}_2 + \mathscr{E}_3 + \ldots + \mu' x' - \mu x - \tfrac{1}{2} \{2\mathscr{E}_2 + 3\mathscr{E}_3 + \ldots\} - \text{higher order terms}$$

$$= \mathscr{E}_0 + \mu' x' - \mu x + \text{terms of at least the third order in } M, N, M', N',$$

and the constant terms in this expression represent the path between the conjugate points of the chief ray. The order to which the difference of path vanishes is thus the same as was found for a single surface.

In the equations involving x and x' we may neglect the difference between L and L' and unity. The equations for x and x', on the neglect of the third and higher order terms of $\mathscr{E}$, are then homogeneous of the second order. It follows that we shall find the same values of x and x' if M, N, M', N' are all multiplied by the same factor. The results we have found by considering the ray M, N, M', N' thus hold for a fan of rays which have their points of nearest approach to the chief ray collinear.

The general expression for $\mathscr{E}$ corresponds exactly to that employed when the properties of astigmatic lenses were investigated*. It follows that the theory for the one is directly applicable to the other. In particular the laws for calculating the new coefficients when two systems are combined to form a more elaborate instrument will be the same in both cases. We deduce at once that the coefficients defining the properties of the pencil surrounding a skew ray may be calculated by multiplying together matrices of the fourth order, the elements of which will be known when the chief ray has been traced through the system. From the product matrix all the properties of the pencil if the ray point of view is adopted, or those of the wave if the wave theory is preferred, are immediately determinable.

Part II. Fourth Order Matrices for Skew Pencils

Construction of Standard Matrices for Skew Pencils

Before the properties of any given skew pencil can be calculated it is necessary to construct the matrices for refraction at each surface of the system. The properties of the complete instrument in so far as rays near the known chief ray are concerned are found from the product, in their proper order, of these elementary matrices. The general character of the matrices involved has been described elsewhere†. It will suffice here to say that the matrices to be multiplied may be represented by

$$\begin{pmatrix} p & s \\ r & q \end{pmatrix},$$

where p, q, r, s are themselves square matrices. The compound matrices have the property that the conjugate reciprocal is always of the form

$$\begin{pmatrix} q' & -s' \\ -r' & p' \end{pmatrix},$$

where the accents indicate that the rows and columns of the matrix denoted by the particular letter used have been interchanged. The elements of these matrices are not those which occur directly in such a function as the eikonal, but are very simply

* "On toric lenses," *Trans. Opt. Soc.* **29** 1927–8 (71). "Canonical forms in the theory of asymmetrical optical systems," *Trans. Opt. Soc.* **29** 1927–8 (88). "The primordial coefficients of asymmetrical lenses," *Trans. Opt. Soc.* **29** 1927–8 (167).

† "A method of evaluating the kinematic coefficients of linked systems," *Proc. Phys. Soc.* **41** 1929 (576).

connected with them. If the eikonal is represented by the symmetrical bilinear form

$$(\alpha\ \beta)\begin{pmatrix} u & -w' \\ -w & v \end{pmatrix}\begin{pmatrix} \alpha' \\ \beta' \end{pmatrix},$$

p, q, r, s are given by

$$p = uW', \quad q = W'v, \quad r = W', \quad s = uW'v - w',$$

where W' is the conjugate reciprocal of w.

Application

The significant quantities in skew refraction will be best appreciated by considering a simple example. Suppose the refracting surface is

$$2x = Ay^2 + 2Byz + Cz^2,$$

and that the chief ray is refracted at the origin in the plane $z = 0$, making angles ϕ and ϕ' respectively with the normal. The eikonal is*

$$\mathscr{E} = \frac{C\mathscr{M}^2 - 2B\mathscr{M}\mathscr{N} + A\mathscr{N}^2}{2(AC - B^2)\mathscr{L}},$$

where $\mathscr{L} = \mu'L' - \mu L$, $\mathscr{M} = \mu'M' - \mu M$, $\mathscr{N} = \mu'N' - \mu N$ and (L, M, N) (L', M', N') are the direction cosines of the incident and refracted rays with respect to the axes of x, y, z. We require to replace these direction cosines by direction cosines referred to new rectangular axes of which one coincides with the chief ray itself. Accordingly we substitute

$$\begin{aligned} L &= l\cos\phi - m\sin\phi, & L' &= l'\cos\phi' - m'\sin\phi', \\ M &= l\sin\phi + m\cos\phi, & M' &= l'\sin\phi' + m'\cos\phi', \\ N &= n, & N' &= n'. \end{aligned}$$

Then, recognising that for the near rays we are considering the departures of l and l' from unity are of too high an order to affect the result, we obtain the significant terms

$$\mathscr{E} = \frac{C(\mu'm'\cos\phi' - \mu m\cos\phi)^2 - 2B(\mu'm'\cos\phi' - \mu m\cos\phi)(\mu'n' - \mu n) + A(\mu'n' - \mu n)^2}{2(AC - B^2)(\mu'\cos\phi' - \mu\cos\phi)},$$

or

$$2\mathscr{E} = (\mu m\ \mu n\ \mu'm'\ \mu'n') \times$$

$$\begin{pmatrix} \frac{C\cos^2\phi}{D} & -\frac{B\cos\phi}{D} & -\frac{C\cos\phi\cos\phi'}{D} & \frac{B\cos\phi}{D} \\ -\frac{B\cos\phi}{D} & \frac{A}{D} & \frac{B\cos\phi'}{D} & -\frac{A}{D} \\ -\frac{C\cos\phi\cos\phi'}{D} & \frac{B\cos\phi'}{D} & \frac{C\cos^2\phi'}{D} & -\frac{B\cos\phi'}{D} \\ \frac{B\cos\phi}{D} & -\frac{A}{D} & -\frac{B\cos\phi'}{D} & \frac{A}{D} \end{pmatrix} \begin{pmatrix} \mu m \\ \mu n \\ \mu'm' \\ \mu'n' \end{pmatrix},$$

where $D = (AC - B^2)\delta = (AC - B^2)(\mu'\cos\phi' - \mu\cos\phi)$.

* *Trans. Opt. Soc.* **28** 1926–7 (266).

The matrix factor for this surface is therefore by the rule just quoted

$$\begin{pmatrix} \frac{\cos\phi}{\cos\phi'} & . & . & . \\ . & 1 & . & . \\ \frac{A\delta}{\cos\phi\cos\phi'} & \frac{B\delta}{\cos\phi} & \frac{\cos\phi'}{\cos\phi} & . \\ \frac{B\delta}{\cos\phi'} & C\delta & . & 1 \end{pmatrix}.$$

The only elements of this matrix which are not indicated by the ordinary formulae for primary and secondary conjugate points are those with B as a factor. When the surface is spherical these elements become zero. It is worth noting that the curvatures A, B, C of the surface appear only as simple factors in the matrix, instead of entering in comparatively complicated expressions, as in the eikonal.

When the reference point is moved along the chief ray through the distance t in the image space before the next surface is encountered the change is represented by multiplying by the factor

$$\begin{pmatrix} 1 & . & -t/\mu' & . \\ . & 1 & . & -t/\mu' \\ . & . & 1 & . \\ . & . & . & 1 \end{pmatrix}$$

exactly as when the incidence is normal. As a rule after the next refraction the chief ray will not lie in the plane $z = 0$. If it is desired to adhere strictly to matrix factors for refraction of the type already given the axes will have to be changed by rotation about the chief ray. This rotation, through say the angle α, is effected by multiplying by the factor

$$\begin{pmatrix} \cos\alpha & \sin\alpha & . & . \\ -\sin\alpha & \cos\alpha & . & . \\ . & . & \cos\alpha & \sin\alpha \\ . & . & -\sin\alpha & \cos\alpha \end{pmatrix}.$$

It is immaterial whether this factor occurs before or after the factor effecting the change of reference point.

With some methods of tracing the chief ray the exact procedure that we have described will be as convenient as any. For others it may be considered more simple to disregard the changes in the plane of refraction and work direct from the algebraic equation of the surface referred to fixed axes. We shall only be concerned with the shape of a refracting surface in the immediate neighbourhood of the point at which the chief ray is refracted, and it will therefore be sufficient to consider the eikonal for the surface

$$Ax^2 + By^2 + Cz^2 + D + 2Fyz + 2Gzx + 2Hxy + 2Ix + 2Jy + 2Kz = 0.$$

It has been shown* that the equation required is

$$\begin{vmatrix} A & H & G & I & \mathscr{L} \\ H & B & F & J & \mathscr{M} \\ G & F & C & K & \mathscr{N} \\ I & J & K & D & \mathscr{E} \\ \mathscr{L} & \mathscr{M} & \mathscr{N} & \mathscr{E} & . \end{vmatrix} = 0$$

or

$$(\mathscr{L}\,\mathscr{M}\,\mathscr{N}\,\mathscr{E}) \begin{pmatrix} a & h & g & i \\ h & b & f & j \\ g & f & c & k \\ i & j & k & d \end{pmatrix} \begin{pmatrix} \mathscr{L} \\ \mathscr{M} \\ \mathscr{N} \\ \mathscr{E} \end{pmatrix} = 0$$

or

$$2\mathscr{E} = -\frac{a\mathscr{L}^2 + b\mathscr{M}^2 + c\mathscr{N}^2 + 2f\mathscr{M}\mathscr{N} + 2g\mathscr{N}\mathscr{L} + 2h\mathscr{L}\mathscr{M}}{i\mathscr{L} + j\mathscr{M} + k\mathscr{N} + \frac{1}{2}d\mathscr{E}},$$

where the matrix $(a\ b\ c\ d)$ is the reciprocal of $(A\ B\ C\ D)$. We assume that the chief ray is refracted at the origin, so that $D = 0$. This implies

$$\begin{vmatrix} a & h & g \\ h & b & f \\ g & f & c \end{vmatrix} = 0,$$

so that the numerator will require simplification. In fact if we add the suffix $_0$ to the variable to indicate the chief ray, the conditions that this ray is refracted at the origin are

$$a\mathscr{L}_0 + h\mathscr{M}_0 + g\mathscr{N}_0 = 0,$$
$$h\mathscr{L}_0 + b\mathscr{M}_0 + f\mathscr{N}_0 = 0,$$
$$g\mathscr{L}_0 + f\mathscr{M}_0 + c\mathscr{N}_0 = 0.$$

It follows that the numerator may be expressed as a function of

$$\mathscr{L} - \mathscr{L}_0,\ \mathscr{M} - \mathscr{M}_0,\ \mathscr{N} - \mathscr{N}_0,$$

and hence in the denominator, since the numerator is of the second order of small quantities, we may replace $\mathscr{L}, \mathscr{M}, \mathscr{N}$ by $\mathscr{L}_0, \mathscr{M}_0, \mathscr{N}_0$. Moreover we can neglect $\mathscr{E}$ in the denominator, since it only serves to introduce higher order terms than any we are at present interested in. Another way of stating this is to say that instead of representing the surface near the origin by the osculating surface of the second order we may employ the osculating paraboloid, the condition for which is $d = 0$.

Since $D = 0$, the tangent plane at the origin where the chief ray is refracted is

$$Ix + Jy + Kz = 0,$$

and therefore

$$\frac{\mathscr{L}_0}{I} = \frac{\mathscr{M}_0}{J} = \frac{\mathscr{N}_0}{K} = \frac{\mu' \cos\phi' - \mu\cos\phi}{(I^2 + J^2 + K^2)^{\frac{1}{2}}} \equiv \frac{\delta}{E}$$
$$= i\mathscr{L}_0 + j\mathscr{M}_0 + k\mathscr{N}_0$$

as $iI + jJ + kK = 1$. The denominator in the expression for $\mathscr{E}$ is thus derived at once from the tracing of the chief ray.

* *Trans. Opt. Soc.* **28** 1926–7 (269).

It is quite a simple matter to simplify the numerator of $\mathscr{E}$ so that the quantities a, h, g, to take a single example, are eliminated. This leads directly to expressions which are not of greater interest or convenience than those obtained without this transformation, and we therefore proceed directly from the numerator in its existing form.

We require in the first place to express the direction of any ray of the pencil in terms of variables which will be small for all members of the pencil. A suitable substitution is

$$\begin{aligned} L &= lL_0 + mL_1 + nL_2, \\ M &= lM_0 + mM_1 + nM_2, \\ N &= lN_0 + mN_1 + nN_2. \end{aligned}$$

The chief ray is thus given by $l = 1, m = n = 0$. Since the numerator is a function of $\mathscr{L} - \mathscr{L}_0$, $\mathscr{M} - \mathscr{M}_0$, $\mathscr{N} - \mathscr{N}_0$, l and l' will only enter in the forms $l - 1$ and $l' - 1$. If we require the condition $l^2 + m^2 + n^2 = 1$ to be satisfied identically, $l - 1$ will be negligibly small compared with m and n, and we may at once neglect the terms with suffix $_0$ when substituting for L, M, N in the numerator.

Let us write

$$\begin{pmatrix} a & h & g \\ h & b & f \\ g & f & c \end{pmatrix} = \mathscr{D},$$

$$\begin{vmatrix} A & H & G & I \\ H & B & F & J \\ G & F & C & K \\ I & J & K & . \end{vmatrix} = \Delta,$$

and denote by 1 either of the linear matrices

$$(L_1\ M_1\ N_1) \quad \text{or} \quad \begin{pmatrix} L_1 \\ M_1 \\ N_1 \end{pmatrix},$$

and similarly by $2'$ either

$$(L_2'\ M_2'\ N_2') \quad \text{or} \quad \begin{pmatrix} L_2' \\ M_2' \\ N_2' \end{pmatrix},$$

the particular matrix meant being obvious from the position. We also put

$$E = (I^2 + J^2 + K^2)^{\frac{1}{2}}, \quad \delta = \mu' \cos \phi' - \mu \cos \phi,$$

and denote the direction of the normal to the surface at the origin by ϵ. Then we may evidently write

$$2\mathscr{E} = -\frac{E}{\delta}(\mu m\ \mu n\ \mu' m'\ \mu' n') \begin{pmatrix} 1\mathscr{D}1 & 1\mathscr{D}2 & -1\mathscr{D}1' & -1\mathscr{D}2' \\ 2\mathscr{D}1 & 2\mathscr{D}2 & -2\mathscr{D}1' & -2\mathscr{D}2' \\ -1'\mathscr{D}1 & -1'\mathscr{D}2 & 1'\mathscr{D}1' & 1'\mathscr{D}2' \\ -2'\mathscr{D}1 & -2'\mathscr{D}2 & 2'\mathscr{D}1' & 2'\mathscr{D}2' \end{pmatrix} \begin{pmatrix} \mu m \\ \mu n \\ \mu' m' \\ \mu' n' \end{pmatrix}.$$

Now from the identities

$$\frac{I^2}{bc-f^2}=\frac{J^2}{ca-g^2}=\frac{K^2}{ab-h^2}=\frac{JK}{gh-fa}=\frac{KI}{hf-gb}=\frac{IJ}{fg-hc}=-\Delta$$

it readily follows that

$$\pi\mathscr{D}\rho\,.\,\sigma\mathscr{D}\tau-\pi\mathscr{D}\tau\,.\,\sigma\mathscr{D}\rho=-\frac{E^2}{\Delta}\,|\,\pi\epsilon\sigma\,|\,|\,\rho\epsilon\tau\,|\,,$$

where π, ρ, σ, τ are linear matrices of the same form as 1, and the determinants on the right are constructed from the direction cosines of the three directions named: for example,

$$E\,|\,2\epsilon1'\,|=\begin{vmatrix} L_2 & M_2 & N_2 \\ I & J & K \\ L_1' & M_1' & N_1' \end{vmatrix}.$$

It is now easy to see that the standard matrix corresponding to this refraction is

$$\begin{pmatrix} \dfrac{|\,1\epsilon2'\,|}{|\,1'\epsilon2'\,|} & -\dfrac{|\,1\epsilon1'\,|}{|\,1'\epsilon2'\,|} & . & . \\ \dfrac{|\,2\epsilon2'\,|}{|\,1'\epsilon2'\,|} & -\dfrac{|\,2\epsilon1'\,|}{|\,1'\epsilon2'\,|} & . & . \\ \mathscr{H}.2\mathscr{D}2' & -\mathscr{H}.2\mathscr{D}1' & -\dfrac{|\,2\epsilon1'\,|}{|\,1\epsilon2\,|} & -\dfrac{|\,2\epsilon2'\,|}{|\,1\epsilon2\,|} \\ -\mathscr{H}.1\mathscr{D}2' & \mathscr{H}.1\mathscr{D}1' & \dfrac{|\,1\epsilon1'\,|}{|\,1\epsilon2\,|} & \dfrac{|\,1\epsilon2'\,|}{|\,1\epsilon2\,|} \end{pmatrix},$$

where

$$\mathscr{H}=\frac{\delta\Delta}{E^3\,|\,1\epsilon2\,|\,|\,1'\epsilon2'\,|}.$$

Since Δ occurs as a factor in $\mathscr{H}$, it is unnecessary to evaluate this determinant, for its product with the elements of $\mathscr{D}$ merely gives the first minors of Δ. No exceptional treatment would therefore be required for such a case as $\Delta = 0$. The six coefficients A, B, C, F, G, H enter linearly into the standard matrix, and the factor E^{-3} in $\mathscr{H}$ serves to make the matrix elements dependent only on the mutual ratios of $A, B, C, F, G, H, I, J, K$.

It is of interest to observe that the direction cosines of the chief ray have so far only appeared explicitly in the factor δ. No use has been made of any relations connecting $L_0, M_0, N_0, L_1, M_1, N_1, L_2, M_2, N_2$. This does not mean that we may take them to define oblique axes, for the argument which led to the neglect of the terms $l-1$ rests upon the identical relation $l^2+m^2+n^2=1$, and this is only identically satisfied when the three temporary reference axes are orthogonal. Without loss of generality we may suppose the signs of $L_1, M_1, N_1, L_2, M_2, N_2$ to be so chosen that

$$L_0=M_1N_2-M_2N_1,\ M_0=N_1L_2-N_2L_1,\ N_0=L_1M_2-L_2M_1.$$

We shall then have $|\,1\epsilon2\,|+\cos\phi=|\,1'\epsilon2'\,|+\cos\phi'=0$.

When the formula given above for the standard matrix is employed in routine calculations, the additional reference directions should be chosen by some invariable rule, such as

$$\begin{pmatrix} L_1 & M_1 & N_1 \\ L_2 & M_2 & N_2 \end{pmatrix} = \begin{pmatrix} -M_0 & 1-\dfrac{M_0^2}{1+L_0} & -\dfrac{M_0N_0}{1+L_0} \\ -N_0 & -\dfrac{M_0N_0}{1+L_0} & 1-\dfrac{N_0^2}{1+L_0} \end{pmatrix}.$$

The particular choice made should depend on the variables used in tracing the chief ray.

When it is desired to exhibit the influence of the coefficients A, B, C, F, G, H explicitly we may modify the matrix by using the identity

$$(L\,M\,N)\,\mathscr{D}\begin{pmatrix} L' \\ M' \\ N' \end{pmatrix} + (\alpha\,\beta\,\gamma)\begin{pmatrix} A & H & G \\ H & B & F \\ G & F & C \end{pmatrix}\begin{pmatrix} \alpha' \\ \beta' \\ \gamma' \end{pmatrix} = 0,$$

where

$$\alpha = MK - NJ, \quad \beta = NI - LK, \quad \gamma = LJ - MI,$$
$$\alpha' = M'K - N'J, \quad \beta' = N'I - L'K, \quad \gamma' = L'J - M'I.$$

The Significance of the Matrix Elements

When the matrix product for the complete optical train has been obtained, the significance of the various elements has to be considered. The forms obtained for these in the course of the work leads us to interpret the quantities involved most readily by regarding I, J, K as pure numbers, and consequently A, B, C, F, G, H as curvatures. If the matrix is represented by

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix}$$

the eight elements $a_{11}, a_{12}, a_{21}, a_{22}, a_{33}, a_{34}, a_{43}, a_{44}$ are ratios, the four $a_{31}, a_{32}, a_{41}, a_{42}$ represent powers since they are of the dimensions of a curvature, and the remainder $a_{13}, a_{14}, a_{23}, a_{24}$ represent distances as they are of the dimensions of a length. There are six independent relations between these sixteen quantities, for the conjugate reciprocal of the matrix is

$$\begin{pmatrix} a_{33} & a_{43} & -a_{13} & -a_{23} \\ a_{34} & a_{44} & -a_{14} & -a_{24} \\ -a_{31} & -a_{41} & a_{11} & a_{21} \\ -a_{32} & -a_{42} & a_{12} & a_{22} \end{pmatrix}.$$

In the product of these two matrices four terms vanish identically, and the remainder consist of six terms each appearing twice. The number of degrees of freedom is therefore ten, the number of terms of the second order in the eikonal or the characteristic function.

The coefficients of the eikonal when the external parts of the chief ray are taken as axes are readily found. The conjugate reciprocal of

$$\begin{pmatrix} a_{31} & a_{32} \\ a_{41} & a_{42} \end{pmatrix} \text{ is } -\begin{pmatrix} e_{31} & e_{32} \\ e_{41} & e_{42} \end{pmatrix}$$

in the notation used earlier in this paper, and

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} e_{31} & e_{32} \\ e_{41} & e_{42} \end{pmatrix} = -\begin{pmatrix} e_{11} & e_{12} \\ e_{21} & e_{22} \end{pmatrix},$$

$$\begin{pmatrix} e_{31} & e_{32} \\ e_{41} & e_{42} \end{pmatrix}\begin{pmatrix} a_{33} & a_{34} \\ a_{43} & a_{44} \end{pmatrix} = -\begin{pmatrix} e_{33} & e_{34} \\ e_{43} & e_{44} \end{pmatrix}.$$

The physical or geometrical significance of the terms or of any relation between them is usually most readily appreciated by utilising the properties of the eikonal or of the characteristic function. The conditions which it is usually desired to fulfil in an optical system can be stated readily in terms of the matrix elements. For example, the conditions that the reference points on the chief ray should be conjugate points entirely free from astigmatism are $a_{13} = a_{23} = a_{14} = a_{24} = 0$. These four conditions are not independent, for the relations

$$a_{11}a_{23} + a_{12}a_{24} = a_{13}a_{21} + a_{14}a_{22}$$

and

$$a_{14}a_{33} + a_{24}a_{43} = a_{34}a_{13} + a_{44}a_{23}$$

are satisfied identically. It follows that all four vanish when three become zero. The number of conditions for freedom from astigmatism at any point is thus three, in agreement with a conclusion reached in an earlier investigation.

The deduction of any characteristics of the system from the properties of the eikonal or of the characteristic function involves no principles which have not been described at some length in other papers printed in these *Transactions*, and they are therefore not repeated here.

Part III. Sixth Order Matrices for Skew Pencils

Observations on Fourth Order Matrices

The expression of the properties of a skew pencil by means of a fourth order matrix is not entirely in keeping with our sense of mathematical fitness. When the imagery is linear in character, for example imagery about the axis of a symmetrical instrument, we obtain a matrix of the second order. When we consider imagery about a ray which always lies in a plane, and is thus specified by two coordinates, we obtain a matrix of the fourth order. By analogy we should expect imagery about a ray which ventures into three dimensional space to involve a matrix of the sixth order. If we review the procedure that has been followed we shall realise that we have been led to a matrix of the fourth order because the three direction cosines of any line are connected by an invariant relation which reduces the number of degrees of freedom from three to two. A treatment leading to matrices of the sixth order

could doubtless be constructed on the assumption that if we maintain homogeneity in the direction cosines throughout our work we may disregard this connection and treat the variables as though they could be varied independently, but it is doubtful if such a course would prove as convincing as a treatment in which difficulties of this type are avoided.

The view that the existence of only two degrees of freedom has forced the expressions we have obtained to assume an unnatural shape receives immediate support from the expressions themselves. The fact that the elements are determinants produces an uncomfortable appearance in the matrices. Our dissatisfaction is further increased when we realise that the quantities $L_1, M_1, N_1, L_2, M_2, N_2$ for any but the extreme sections of a ray are meaningless. They involve arbitrary factors which are entirely foreign to the problem. The only direction which can possibly be significant for this part of the ray is the direction of the ray itself, and the quantities L_0, M_0, N_0 which specify it are absent. In fact we cannot appreciate the grace of the edifice we are erecting because the scaffolding still remains. We therefore proceed to consider how the matrices can be purged of all irrelevant factors. It will be found that this leads naturally to the expression of the properties of the pencil by a matrix of the sixth order.

Relations between Reference Directions

The direction cosines with suffixes 1 and 2 are evidently needed for the external parts of the ray. They only become objectionable when they relate to an internal part of the optical path, that is to say, when they persist after the matrices have been multiplied together. We can of course be certain in advance that they invariably combine in such a way that only relevant factors remain. In other words the products are expressible in terms of L_0, M_0, N_0 only. The available relations are

$$L_0 = M_1N_2 - N_1M_2, \qquad M_0 = N_1L_2 - L_1N_2, \qquad N_0 = L_1M_2 - M_1L_2,$$

$$M_0^2 + N_0^2 = L_1^2 + L_2^2, \qquad N_0^2 + L_0^2 = M_1^2 + M_2^2, \qquad L_0^2 + M_0^2 = N_1^2 + N_2^2,$$

$$-M_0N_0 = M_1N_1 + M_2N_2, \quad -N_0L_0 = N_1L_1 + N_2L_2, \quad -L_0M_0 = L_1M_1 + L_2M_2.$$

We continue as before to designate either of the linear matrices $(L_\sigma\ M_\sigma\ N_\sigma)$, $\begin{pmatrix} L_\sigma \\ M_\sigma \\ N_\sigma \end{pmatrix}$ by the suffix σ alone, and add the convention that the skew matrix

$$\begin{pmatrix} . & -N_\sigma & M_\sigma \\ N_\sigma & . & -L_\sigma \\ -M_\sigma & L_\sigma & . \end{pmatrix}$$

may be denoted by (σ). We will avoid the use of the numeral o to denote the direction of the chief ray, and write η and (η) in place of o and (o) respectively. As before, ϵ will relate to the normal to the refracting surface at the point where the chief ray is incident.

When the products involve a vertical matrix on the left and a horizontal matrix on the right, we have

$$2.1 - 1.2 = \begin{pmatrix} L_2 \\ M_2 \\ N_2 \end{pmatrix} (L_1\, M_1\, N_1) - \begin{pmatrix} L_1 \\ M_1 \\ N_1 \end{pmatrix} (L_2\, M_2\, N_2)$$

$$= \begin{pmatrix} . & L_2M_1 - L_1M_2 & L_2N_1 - L_1N_2 \\ M_2L_1 - M_1L_2 & . & M_2N_1 - M_1N_2 \\ N_2L_1 - N_1L_2 & N_2M_1 - N_1M_2 & . \end{pmatrix}$$

$$= \begin{pmatrix} . & -N_0 & M_0 \\ N_0 & . & -L_0 \\ -M_0 & L_0 & . \end{pmatrix} = (\eta)$$

and

$$1.1 + 2.2 = \begin{pmatrix} L_1 \\ M_1 \\ N_1 \end{pmatrix} (L_1\, M_1\, N_1) + \begin{pmatrix} L_2 \\ M_2 \\ N_2 \end{pmatrix} (L_2\, M_2\, N_2)$$

$$= \begin{pmatrix} L_1^2 + L_2^2 & L_1M_1 + L_2M_2 & L_1N_1 + L_2N_2 \\ M_1L_1 + M_2L_2 & M_1^2 + M_2^2 & M_1N_1 + M_2N_2 \\ N_1L_1 + N_2L_2 & N_1M_1 + N_2M_2 & N_1^2 + N_2^2 \end{pmatrix}$$

$$= \begin{pmatrix} M_0^2 + N_0^2 & -L_0M_0 & -L_0N_0 \\ -L_0M_0 & L_0^2 + N_0^2 & -M_0N_0 \\ -L_0N_0 & -M_0N_0 & L_0^2 + M_0^2 \end{pmatrix} = -(\eta)^2.$$

Derivation of Sixth Order Matrices

We now observe that all the elements of the fourth order matrix for refraction at a single surface may be written as products in which the leading factor is either 1 or 2 and the final factor either 1′ or 2′. This is accomplished by applying the identity $|\pi\epsilon\rho| = \pi(\epsilon)\rho$ to the numerators of the eight elements written as fractions. The matrix now takes a form included in the general type

$$\begin{pmatrix} 1\mathfrak{p}2' & -1\mathfrak{p}1' & -1\mathfrak{s}1' & -1\mathfrak{s}2' \\ 2\mathfrak{p}2' & -2\mathfrak{p}1' & -2\mathfrak{s}1' & -2\mathfrak{s}2' \\ 2\mathfrak{r}2' & -2\mathfrak{r}1' & -2\mathfrak{q}1' & -2\mathfrak{q}2' \\ -1\mathfrak{r}2' & 1\mathfrak{r}1' & 1\mathfrak{q}1' & 1\mathfrak{q}2' \end{pmatrix},$$

where the numerical characters denote linear matrices and the literal characters square matrices. This expression may be expanded into the form

$$\begin{pmatrix} 1 & . & . & . \\ . & 2 & . & . \\ . & . & 2 & . \\ . & . & . & -1' \end{pmatrix} \Phi \begin{pmatrix} 2' & . & . & . \\ . & -1' & . & . \\ . & . & -1' & . \\ . & . & . & -2' \end{pmatrix},$$

where
$$\Phi = \begin{pmatrix} \mathrm{p} & \mathrm{p} & \mathrm{s} & \mathrm{s} \\ \mathrm{p} & \mathrm{p} & \mathrm{s} & \mathrm{s} \\ \mathrm{r} & \mathrm{r} & \mathrm{q} & \mathrm{q} \\ \mathrm{r} & \mathrm{r} & \mathrm{q} & \mathrm{q} \end{pmatrix}.$$

When we combine matrices to represent a first refraction, a translation, and a second refraction we therefore have, apart from the extreme factors, the product

$$\Phi \begin{pmatrix} 2' & . & . & . \\ . & -1' & . & . \\ . & . & -1' & . \\ . & . & . & -2' \end{pmatrix} T \begin{pmatrix} 1' & . & . & . \\ . & 2' & . & . \\ . & . & 2' & . \\ . & . & . & -1' \end{pmatrix} \Phi',$$

where
$$T = \begin{pmatrix} 1 & . & -t/\mu' & . \\ . & 1 & . & -t/\mu' \\ . & . & 1 & . \\ . & . & . & 1 \end{pmatrix},$$

and the character 1 in the diagonal of this matrix represents unity. The product of the three central matrices yields

$$\begin{pmatrix} 2'1' & . & -2'2't/\mu' & . \\ . & -1'2' & . & -1'1't/\mu' \\ . & . & -1'2' & . \\ . & . & . & 2'1' \end{pmatrix},$$

and on multiplying before and behind by Φ and Φ' respectively we obtain

$$\bar{\Phi} = \begin{pmatrix} \bar{\mathrm{p}} & \bar{\mathrm{p}} & \bar{\mathrm{s}} & \bar{\mathrm{s}} \\ \bar{\mathrm{p}} & \bar{\mathrm{p}} & \bar{\mathrm{s}} & \bar{\mathrm{s}} \\ \bar{\mathrm{r}} & \bar{\mathrm{r}} & \bar{\mathrm{q}} & \bar{\mathrm{q}} \\ \bar{\mathrm{r}} & \bar{\mathrm{r}} & \bar{\mathrm{q}} & \bar{\mathrm{q}} \end{pmatrix},$$

where
$$\begin{pmatrix} \bar{\mathrm{p}} & \bar{\mathrm{s}} \\ \bar{\mathrm{r}} & \bar{\mathrm{q}} \end{pmatrix} = \begin{pmatrix} \mathrm{p} & \mathrm{s} \\ \mathrm{r} & \mathrm{q} \end{pmatrix} \begin{pmatrix} (\eta') & (\eta')^2 t/\mu' \\ . & (\eta') \end{pmatrix} \begin{pmatrix} \mathrm{p}' & \mathrm{s}' \\ \mathrm{r}' & \mathrm{q}' \end{pmatrix}.$$

It follows that the properties of the system are given by multiplying together sixth order matrices of the type

$$\begin{pmatrix} \mathrm{p} & . \\ \mathrm{r} & \mathrm{q} \end{pmatrix}$$

for a single refraction and of the type

$$\begin{pmatrix} (\eta) & (\eta)^2 t/\mu \\ . & (\eta) \end{pmatrix}$$

for the medium separating the surfaces. It should be particularly noted that this matrix no longer signifies a change of origin only. It does not degenerate to a unit matrix when successive surfaces touch one another. The essential similarity between

the two types of matrix involved will be more evident if $(\eta)^2 t/\mu$ is written in the form $(\eta)\,\frac{t}{\mu}\,(\eta)$, and if we note that $\frac{\mathscr{D}}{E^2} = (\epsilon)\,d\,(\epsilon)$, where d denotes

$$\begin{pmatrix} A & H & G \\ H & B & F \\ G & F & C \end{pmatrix},$$

the matrix of the second order coefficients in the equation of the surface.

The Matrix and the Eikonal

By multiplying together the matrices for the surfaces and internal media of the system we obtain the four third order matrices, $\mathbf{p}, \mathbf{q}, \mathbf{r}, \mathbf{s}$ in which all the properties of the pencil are involved. The corresponding fourth order matrix is

$$\begin{pmatrix} 1\mathbf{p}2' & -1\mathbf{p}1' & -1\mathbf{s}1' & -1\mathbf{s}2' \\ 2\mathbf{p}2' & -2\mathbf{p}1' & -2\mathbf{s}1' & -2\mathbf{s}2' \\ 2\mathbf{r}2' & -2\mathbf{r}1' & -2\mathbf{q}1' & -2\mathbf{q}2' \\ -1\mathbf{r}2' & 1\mathbf{r}1' & 1\mathbf{q}1' & 1\mathbf{q}2' \end{pmatrix},$$

where 1 and 2 now relate to the initial, and 1′ and 2′ to the final portions of the chief ray. The elements of this matrix are ordinary numbers, and the first step in establishing the eikonal is to obtain the conjugate reciprocal of

$$\begin{pmatrix} 2\mathbf{r}2' & -2\mathbf{r}1' \\ -1\mathbf{r}2' & 1\mathbf{r}1' \end{pmatrix}.$$

It is easily verified that the determinant of this matrix is $\eta\mathscr{R}\eta'$, where $\mathscr{R}$ is the matrix of the co-factors of $\mathbf{r}$ in $|\mathbf{r}|$.

The expression for the eikonal then readily follows in the form

$$2\eta\mathscr{R}\eta'\mathscr{E} = (\mu m\ \mu n\ \mu' m'\ \mu' n')$$

$$\begin{pmatrix} 1\mathbf{p}(\eta')\mathbf{r}'1 & 1\mathbf{p}(\eta')\mathbf{r}'2 & -1\mathbf{r}1' & -1\mathbf{r}2' \\ 2\mathbf{p}(\eta')\mathbf{r}'1 & 2\mathbf{p}(\eta')\mathbf{r}'2 & -2\mathbf{r}1' & -2\mathbf{r}2' \\ -1'\mathbf{r}'1 & -1'\mathbf{r}'2 & 1'\mathbf{r}'(\eta)\mathbf{q}1' & 1'\mathbf{r}'(\eta)\mathbf{q}2' \\ -2'\mathbf{r}'1 & -2'\mathbf{r}'2 & 2'\mathbf{r}'(\eta)\mathbf{q}1' & 2'\mathbf{r}'(\eta)\mathbf{q}2' \end{pmatrix} \begin{pmatrix} \mu m \\ \mu n \\ \mu' m' \\ \mu' n \end{pmatrix}$$

$$= (\overline{\mu m 1 + n 2}\ \ \mu' \overline{m'1' + n'2'}) \begin{pmatrix} \mathbf{p}(\eta')\mathbf{r}' & -\mathbf{r} \\ -\mathbf{r}' & \mathbf{r}'(\eta)\mathbf{q} \end{pmatrix} \begin{pmatrix} \overline{\mu m 1 + n 2} \\ \mu' \overline{m'1' + n'2'} \end{pmatrix},$$

where $\mathbf{r}'$ is the conjugate of $\mathbf{r}$. Now from the definitions of m, n, m', n', recalling that l and l' differ insensibly from unity, we have

$$(m1 + n2) = (L - L_0 \quad M - M_0 \quad N - N_0),$$
$$(m'1' + n'2') = (L' - L_0' \quad M' - M_0' \quad N' - N_0'),$$

and the sixth order matrix thus yields the coefficients of the eikonal referred to the original axes. In exactly the same way, by forming the conjugate reciprocal of the portion of the fourth order matrix which involves $\mathbf{s}$ we may find the coefficients of the characteristic function.

The Identities

The matrix of the sixth order contains 36 elements, and the general quadratic with six variables has 21 terms. We should therefore expect to find in general 15 identical relations satisfied by the elements of the matrix. The identities of the fourth order matrix take the form that

$$\begin{pmatrix} 1p2' & -1p1' & -1s1' & -1s2' \\ 2p2' & -2p1' & -2s1' & -2s2' \\ 2r2' & -2r1' & -2q1' & -2q2' \\ -1r2' & 1r1' & 1q1' & 1q2' \end{pmatrix} \text{ and } \begin{pmatrix} -2q1' & 1q1' & 1s1' & 2s1' \\ -2q2' & 1q2' & 1s2' & 2s2' \\ -2r2' & 1r2' & 1p2' & 2p2' \\ 2r1' & -1r1' & -1p1' & -2p1' \end{pmatrix}$$

are conjugate reciprocals of one another. It follows almost immediately from the product in one order that

$$\begin{pmatrix} (\eta) & . \\ . & (\eta) \end{pmatrix} \begin{pmatrix} \mathbf{p} & \mathbf{s} \\ \mathbf{r} & \mathbf{q} \end{pmatrix} \begin{pmatrix} (\eta') & . \\ . & (\eta') \end{pmatrix} \begin{pmatrix} \mathbf{q}' & \mathbf{s}' \\ \mathbf{r}' & \mathbf{p}' \end{pmatrix} \begin{pmatrix} (\eta) & . \\ . & (\eta) \end{pmatrix} + \begin{pmatrix} (\eta) & . \\ . & (\eta) \end{pmatrix} = 0,$$

and in the reverse order that

$$\begin{pmatrix} (\eta') & . \\ . & (\eta') \end{pmatrix} \begin{pmatrix} \mathbf{q}' & \mathbf{s}' \\ \mathbf{r}' & \mathbf{p}' \end{pmatrix} \begin{pmatrix} (\eta) & . \\ . & (\eta) \end{pmatrix} \begin{pmatrix} \mathbf{p} & \mathbf{s} \\ \mathbf{r} & \mathbf{q} \end{pmatrix} \begin{pmatrix} (\eta') & . \\ . & (\eta') \end{pmatrix} + \begin{pmatrix} (\eta') & . \\ . & (\eta') \end{pmatrix} = 0,$$

where $\mathbf{p}', \mathbf{q}', \mathbf{r}', \mathbf{s}'$ are the conjugates of $\mathbf{p}, \mathbf{q}, \mathbf{r}, \mathbf{s}$.

An alternative form for these equations, obtained by introducing the identity $(\eta)^3 + (\eta) = 0$, is, for the first set,

$$(\eta)\,[\mathbf{p}\,(\eta')\,\mathbf{q}' + \mathbf{s}\,(\eta')\,\mathbf{r}' - (\eta)]\,(\eta) = 0 \qquad (\eta)\,[\mathbf{p}\,(\eta')\,\mathbf{s}' + \mathbf{s}\,(\eta')\,\mathbf{p}']\,(\eta) = 0$$
$$(\eta)\,[\mathbf{r}\,(\eta')\,\mathbf{q}' + \mathbf{q}\,(\eta')\,\mathbf{r}']\,(\eta) = 0 \qquad (\eta)\,[\mathbf{r}\,(\eta')\,\mathbf{s}' + \mathbf{q}\,(\eta')\,\mathbf{p}' - (\eta)]\,(\eta) = 0,$$

and for the second set

$$(\eta')\,[\mathbf{q}'\,(\eta)\,\mathbf{p} + \mathbf{s}'\,(\eta)\,\mathbf{r} - (\eta')]\,(\eta') = 0 \qquad (\eta')\,[\mathbf{q}'\,(\eta)\,\mathbf{s} + \mathbf{s}'\,(\eta)\,\mathbf{q}]\,(\eta') = 0$$
$$(\eta')\,[\mathbf{r}'\,(\eta)\,\mathbf{p} + \mathbf{p}'\,(\eta)\,\mathbf{r}]\,(\eta') = 0 \qquad (\eta')\,[\mathbf{r}'\,(\eta)\,\mathbf{s} + \mathbf{p}'\,(\eta)\,\mathbf{q} - (\eta')]\,(\eta') = 0.$$

It should be realised that we may not, without special justification, remove the end factors (η) and (η') from these equations, since both denote singular matrices. An equation of the form $(\eta)\,(\sigma)\,(\eta) = 0$ implies that σ is a skew-symmetrical matrix, and that $\eta\sigma = 0$. Now since the conjugate of (η) is $-(\eta)$, an expression such as $\mathbf{p}\,(\eta')\,\mathbf{s}' + \mathbf{s}\,(\eta')\,\mathbf{p}'$ is skew symmetrical in virtue of its construction. Such an equation therefore leads only to a single condition. On the other hand such an equation as $(\eta)\,[\mathbf{p}\,(\eta')\,\mathbf{q}' + \mathbf{s}\,(\eta')\,\mathbf{r}' - (\eta)]\,(\eta) = 0$ involves three conditions for the vanishing of the diagonal elements, three conditions for the correspondence of non-diagonal elements, and one condition of the form $\eta\sigma = 0$, making a total of seven conditions. The fourth relation of the set introduces no new condition—it is clearly the conjugate of the first relation. From the first set we therefore obtain $7 + 1 + 1 = 9$ relations. Six additional independent relations thus remain to be found. The two sets of relations given for the fourth order matrices are equivalent to one another, and if one set is given the other can be deduced provided any one of the second order matrices obtained by taking a corner element and the three next to it is non-

singular. We cannot apply a parallel method to the sixth order matrix because all the four third order matrices into which the whole matrix could be divided for the purpose are singular. The equivalence of the two sets may however be shown in the following way. From the equations of the first row we have

$$\begin{aligned} 0 &= (\eta)\,[\mathfrak{p}\,(\eta')\,\mathfrak{q}' + \mathfrak{s}\,(\eta')\,\mathfrak{r}' - (\eta)]\,(\eta)\,\mathfrak{p}\,(\eta') + (\eta)\,[\mathfrak{p}\,(\eta')\,\mathfrak{s}' + \mathfrak{s}\,(\eta')\,\mathfrak{p}']\,(\eta)\,(\mathfrak{r})\,(\eta') \\ &= (\eta)\,\mathfrak{p}\,(\eta')\,[\mathfrak{q}'\,(\eta)\,\mathfrak{p} + \mathfrak{s}'\,(\eta)\,\mathfrak{r} - (\eta')]\,(\eta') + (\eta)\,\mathfrak{s}\,(\eta')\,[\mathfrak{r}'\,(\eta)\,\mathfrak{p} + \mathfrak{p}'\,(\eta)\,\mathfrak{r}]\,(\eta'), \end{aligned}$$

since

$$(\eta)^2\,\mathfrak{p}\,(\eta') = -\,(\eta)\,\mathfrak{p}\,(\eta') = (\eta)\,\mathfrak{p}\,(\eta')^2,$$

and similarly from the second row

$$\begin{aligned} 0 &= (\eta)\,[\mathfrak{r}\,(\eta')\,\mathfrak{q}' + \mathfrak{q}\,(\eta')\,\mathfrak{r}']\,(\eta)\,\mathfrak{p}\,(\eta') + (\eta)\,[\mathfrak{r}\,(\eta')\,\mathfrak{s}' + \mathfrak{q}\,(\eta')\,\mathfrak{p}' - (\eta)]\,(\eta)\,\mathfrak{r}\,(\eta') \\ &= (\eta)\,\mathfrak{r}\,(\eta')\,[\mathfrak{q}'\,(\eta)\,\mathfrak{p} + \mathfrak{s}'\,(\eta)\,\mathfrak{r} - (\eta')]\,(\eta') + (\eta)\,\mathfrak{q}\,(\eta')\,[\mathfrak{r}'\,(\eta)\,\mathfrak{p} + \mathfrak{p}'\,(\eta)\,\mathfrak{r}]\,(\eta'). \end{aligned}$$

By considering the simultaneous vanishing of these transformed factors we may deduce the first two relations of the second set. The remaining equations may be considered in a similar way.

We must then look elsewhere for the six additional relations required. Analogy with the fourth order matrix suggests that the expression for the eikonal is quite symmetrical. For example, we have from the coefficients of the two terms in mn

$$\begin{aligned} 2\mathfrak{p}\,(\eta')\,\mathfrak{r}'_1 &= 1\mathfrak{p}\,(\eta')\,\mathfrak{r}'_2 \\ &= -\,2\mathfrak{r}\,(\eta')\,\mathfrak{p}'_1, \end{aligned}$$

and from this we are entitled to deduce

$$(\eta)\,[\mathfrak{p}\,(\eta')\,\mathfrak{r}' + \mathfrak{r}\,(\eta')\,\mathfrak{p}']\,(\eta) = 0,$$

but fuller investigation would be needed to justify the removal of the end factors. In the same way we have

$$\begin{aligned} (\eta)\,[\mathfrak{q}\,(\eta')\,\mathfrak{s}' + \mathfrak{s}\,(\eta')\,\mathfrak{q}']\,(\eta) &= 0, \\ (\eta')\,[\mathfrak{r}'\,(\eta)\,\mathfrak{q} + \mathfrak{q}'\,(\eta)\,\mathfrak{r}]\,(\eta') &= 0, \\ (\eta')\,[\mathfrak{p}'\,(\eta)\,\mathfrak{s} + \mathfrak{s}'\,(\eta)\,\mathfrak{p}]\,(\eta') &= 0. \end{aligned}$$

Even if these were all independent they would only provide four of the six relations we are seeking, and it becomes necessary to examine the equations in greater detail.

It may be shown by considering simple examples that the extreme factors are not in general removable from the relations derived from the fourth order identities. We therefore consider the value of

$$\mathfrak{p}\,(\eta')\,\mathfrak{r}' + \mathfrak{r}\,(\eta')\,\mathfrak{p}'.$$

Let the system be divided into two parts denoted by the suffixes 1 and 2. Taking part 1 to relate to a single surface, we have, by the law of combination,

$$\begin{aligned} &\mathfrak{p}\,(\eta')\,\mathfrak{r}' + \mathfrak{r}\,(\eta')\,\mathfrak{p}' \\ &= -\,\mathfrak{p}_1\,(\eta_1)\,\mathfrak{p}_2\,(\eta_2)\,[\mathfrak{p}_2'\,(\eta_1)\,\mathfrak{r}_1' + \mathfrak{r}_2'\,(\eta_1)\,\mathfrak{q}_1'] - [\mathfrak{r}_1\,(\eta_1)\,\mathfrak{p}_2 + \mathfrak{q}_1\,(\eta_1)\,\mathfrak{r}_2]\,(\eta_2)\,\mathfrak{p}_2'\,(\eta_1)\,\mathfrak{p}_1' \\ &= -\,\mathfrak{p}_1\,(\eta_1)\,\mathfrak{p}_2\,(\eta_2)\,\mathfrak{p}_2'\,(\eta_1)\,\mathfrak{r}_1' - \mathfrak{r}_1\,(\eta_1)\,\mathfrak{p}_2\,(\eta_2)\,\mathfrak{p}_2'\,(\eta_1)\,\mathfrak{p}_1' \\ &\quad -\,\mathfrak{p}_1\,(\eta_1)\,\mathfrak{p}_2\,(\eta_2)\,\mathfrak{r}_2'\,(\eta_1)\,\mathfrak{q}_1' - \mathfrak{q}_1\,(\eta_1)\,\mathfrak{r}_2\,(\eta_2)\,\mathfrak{p}_2'\,(\eta_1)\,\mathfrak{p}_1'. \end{aligned}$$

Now $\mathfrak{p}_2\,(\eta_2)\,\mathfrak{p}_2'$ is skew symmetrical, and therefore $(\eta_1)\,\mathfrak{p}_2\,(\eta_2)\,\mathfrak{p}_2'\,(\eta_1)$ is a constant multiple of (η_1), say $h\,(\eta_1)$.

Also we know that

$$(\eta_1)\,\mathbf{p}_2\,(\eta_2)\,\mathbf{r}_2'\,(\eta_1) = -\,(\eta_1)\,\mathbf{r}_2\,(\eta_2)\,\mathbf{p}_2'\,(\eta_1) = k, \text{ say,}$$

where k is symmetrical. Thus

$$\mathbf{p}\,(\eta')\,\mathbf{r}' + \mathbf{r}\,(\eta')\,\mathbf{p} = -\,h\,[\mathbf{p}_1\,(\eta_1)\,\mathbf{r}_1' + \mathbf{r}_1\,(\eta_1)\,\mathbf{p}_1'] - \mathbf{p}_1 k \mathbf{q}_1' + \mathbf{q}_1 k \mathbf{p}_1'.$$

Now for a single surface the coefficient of h vanishes. Also

$$\mathbf{p}_1 = \frac{(\epsilon_1)}{\cos\phi_1'}, \quad \mathbf{q}_1 = \frac{(\epsilon_1)}{\cos\phi_1}, \quad \mathbf{p}_1' = -\frac{(\epsilon_1)}{\cos\phi_1'}, \quad \mathbf{q}_1' = -\frac{(\epsilon_1)}{\cos\phi_1},$$

so that the remaining terms cancel one another. It follows that

$$\mathbf{p}\,(\eta')\,\mathbf{r}' + \mathbf{r}\,(\eta')\,\mathbf{p}' = 0,$$

however many surfaces there may be in the system. The extreme factors are similarly removable from the three remaining relations. Each of these equations thus yields three conditions instead of one. We must clearly associate two of them with the first set of relations, and two with the other set. The number of relations in each set is then increased from 9 to 15, the number we expected to obtain. We infer that all the relations which are generally true have been found.

The Working Formulae

At this stage it is convenient to repeat the formulae which will be used in numerical calculations, making such alterations in the notation as are convenient.

We assume that the optical system consists of a number of refracting (which of course includes reflecting) surfaces. They may be identified by the serial numbers $1, 2, 3 \ldots, u$, assigned in the order in which a ray meets them. The external media may be numbered 0 and u, the intervening media carrying the numbers $1, 2, 3 \ldots, u-1$. The numbers will usually appear as suffixes to surface constants or to medium or ray constants. The refractive index will be denoted by μ. Any ray is traced through the system. Its successive directions depend only on the refractive indices and the directions of the tangent planes to the surfaces at the points of incidence, that is to say on the first order terms in the equations of the surfaces. The direction cosines of this ray will be denoted by l, m, n, and we shall write

$$\eta \text{ for } (l, m, n),\ \eta' \text{ for } \begin{pmatrix} l \\ m \\ n \end{pmatrix} \text{ and } (\eta) \text{ for } \begin{pmatrix} . & -n & m \\ n & . & -l \\ -m & l & . \end{pmatrix}.$$

When the letter η is replaced by ϵ we signify the corresponding functions of the direction cosines of the normal to the surface. If we denote the angles of incidence and refraction by ϕ and ϕ', we then have

$$\epsilon_v \eta'_{v-1} = \eta_{v-1} \epsilon_v' = \cos\phi_v,$$
$$\epsilon_v \eta_v' = \eta_v \epsilon_v' = \cos\phi_v'.$$

When any ray has been traced we can find all the properties of pencils of rays immediately around it if the second order terms in the equation of each surface are

known, and the distance t_v along this ray between the points of refraction at surfaces v and $v+1$ is known. Let the equation of the surface be

$$Ax^2 + By^2 + Cz^2 + 2Fyz + 2Gzx + 2Hxy + 2Ix + 2Jy + 2Kz = 0,$$

and write $$E = (I^2 + J^2 + K^2)^{\frac{1}{2}}.$$

It will be found that the sign of E determines the sign of the power of the surface. The conventions of all systems concern E and t only. As the applications will normally relate to systems in which one of the quantities I, J, K is numerically large compared with others, the sign of E can be made to agree with that of this particular quantity. For example, if we are dealing with a system in which there is an axis of symmetry parallel to the axis of x, the sign of E will be that of I. The matrix for refraction at surface v is

$$\begin{pmatrix} (\epsilon_v)\sec\phi_v' & . \\ (\epsilon_v)\, d_v\, (\epsilon_v) & (\epsilon_v)\sec\phi_v \end{pmatrix},$$

where $$d_v = \frac{\mu_v \cos\phi_v' - \mu_{v-1}\cos\phi_v}{E_v \cos\phi_v \cos\phi_v'} \begin{pmatrix} A_v & H_v & G_v \\ H_v & B_v & F_v \\ G_v & F_v & C_v \end{pmatrix}.$$

If the reference points are moved away from the point of refraction of the chief ray, we multiply this matrix in front by

$$\begin{pmatrix} \mathrm{I} & -(\eta_{v-1})\rho_{v-1} \\ . & \mathrm{I} \end{pmatrix}$$

for a displacement in the direction in which the chief ray travels through a distance $\rho_{v-1}\mu_{v-1}$ in medium $v-1$, and similarly behind by

$$\begin{pmatrix} \mathrm{I} & (\eta_v)\rho_v \\ . & \mathrm{I} \end{pmatrix}$$

for a corresponding displacement through the distance $\rho_v\mu_v$ in the medium v, where I is the unit third order matrix. When another refraction occurs at the surface $v+1$, we have to introduce the factor

$$\begin{pmatrix} (\epsilon_{v+1})\sec\phi'_{v+1} & . \\ (\epsilon_{v+1})\, d_{v+1}\, (\epsilon_{v+1}) & (\epsilon_{v+1})\sec\phi_{v+1} \end{pmatrix}$$

corresponding to this refraction, but *before we leave the medium v we must introduce the medium matrix* $\begin{pmatrix} (\eta_v) & . \\ . & (\eta_v) \end{pmatrix}$. This is the sole departure from the procedure hitherto followed.

By multiplying the matrices indicated in their proper order, we reach a matrix which may be denoted by

$$\begin{pmatrix} \mathbf{p} & \mathbf{s} \\ \mathbf{r} & \mathbf{q} \end{pmatrix}.$$

This contains all factors except the matrices of the extreme media, i.e. it does not involve (η_0) or (η_u). From this matrix we derive the matrix of eikonal coefficients,

$$\frac{1}{\eta_0 \mathscr{R} \eta'_u} \begin{pmatrix} \mathbf{p}\,(\eta_u)\,\mathbf{r}' & -\mathbf{r} \\ -\mathbf{r}' & \mathbf{r}'\,(\eta_0)\,\mathbf{q} \end{pmatrix},$$

where $\mathscr{R}$ is the matrix of the co-factors of $\mathbf{r}$ in $|\mathbf{r}|$, and $\mathbf{r}'$ is the transposed of $\mathbf{r}$, or the matrix of characteristic function coefficients

$$\frac{1}{\eta_0 \mathscr{S} \eta'_u} \begin{pmatrix} \mathbf{q}(\eta_u)\,\mathbf{s}' & \mathbf{s} \\ \mathbf{s}' & \mathbf{s}'(\eta_0)\,\mathbf{p} \end{pmatrix},$$

where $\mathscr{S}$ is the matrix of the co-factors of $\mathbf{s}$ in $|\mathbf{s}|$, and $\mathbf{s}'$ is the transposed of $\mathbf{s}$.

The variables in the eikonal are

$$\mu_0(L_0 - l_0),\ \mu_0(M_0 - m_0),\ \mu_0(N_0 - n_0),\ \mu_u(L_u - l_u),\ \mu_u(M_u - m_u),\ \mu_u(N_u - n_u)$$

and in the characteristic function are

$$X_0 - x_0,\ Y_0 - y_0,\ Z_0 - z_0,\ X_u - x_u,\ Y_u - y_u,\ Z_u - z_u,$$

where $x_0, y_0, z_0, x_u, y_u, z_u$ are the coordinates of the reference points on the chief ray.

It can easily be verified that these expressions include the better known relations in their simplest form by making suitable substitutions. For example, we obtain the formulae for the pencil about a ray which is undeviated in its passage through the instrument by putting

$$(\epsilon_v) = (\eta_v) = \begin{pmatrix} . & . & . \\ . & . & -1 \\ . & 1 & . \end{pmatrix}$$

for all values of v.

In presenting the paper Mr Smith said: Since this paper was written and accepted I have, thanks to the lending facilities now offered by the Library of the Science Museum, for the first time had an opportunity of reading Prof. R. A. Sampson's "A continuation of Gauss's 'Dioptrische Untersuchungen'" read in 1897 before the London Mathematical Society.* My interest in a later paper† of Prof. Sampson's had made me for a long time desirous of reading this paper also, but I had previously been unable to obtain a copy. After reading the paper I can only express my amazement at its universal neglect. Much that I have written lately in the belief that I was breaking new ground is there dealt with at length. Important corrections in the published histories of Optics in the last thirty years are clearly called for. The paper anticipates much that is now regarded as among the most outstanding original work in optics of this century. I think we can safely assume that the compilers of the best known bibliographies in optics, though they refer to it, have neither read the paper nor realised its importance. I can only express to Prof. Sampson my own apologies for my involuntary failure to make due reference to it in the past, and couple with them the hope that his work will ultimately receive the appreciation which I think it so richly merits.

* *Proc. Lond. Math. Soc.* **29** 1898 (33–83).

† "A New Treatment of Optical Aberrations," *Phil. Trans.* A **212** 1913 (149).

DISCUSSION

Prof. Sampson: Mr Smith has referred to a paper by myself published as long ago as 1897 and has been good enough to say that it ought to have attracted more attention. I am perhaps myself somewhat to blame in the matter, for I did not write about the subject again until 1912 and then in a somewhat different direction. The fact is, I had some doubt whether an analytical method was really the most convenient. Nearly all the great masters have used geometrical methods and have kept closely in touch with their specific problem at every step. Very great precautions are required to keep an algebraic method correct in its applications. Even such a simple matter as reflection is full of pitfalls. Still for wholly general results it has a power that geometrical methods cannot rival. I will mention one point of interest in my 1897 paper, realised after it was written. The analysis there developed is nothing more nor less than the Contact Transformation, which arises in dynamics but is of course applicable to optics. It remains, I believe, the only specific case when the actual coefficients in the Contact Transformation are explicitly known.

Capt. T. Y. Baker: The work which Mr Smith has been doing in the development of a new calculus for optical imagery has been further advanced by this present paper. Unfortunately as in some of the earlier ones the practical value of it to the optical designer is hidden by the general character of the formulae and is not illustrated by any numerical example.

I believe that the use of matrices gives a most useful systematic method of dealing with optical imagery that would be invaluable to practical computers if only Mr Smith would illustrate his papers with numerical cases of the type that the practical man has to deal with, so that the latter, even if he does not understand the complete theory upon which the calculus is based can at least follow the arithmetic. He will at all events be then in a position to judge whether or no these computation methods possess advantages in giving him an organised systematic calculus that can be worked by anyone who learns the rules and doesn't make arithmetical errors.

Without a thorough and complete acquaintance with the earlier papers it is not possible to criticise this paper in detail and I must, therefore, confine myself to a point that I mentioned in the discussion and which I have since been able to go into more fully.

Mr Smith deals, on p. 134, with the case of refraction at a single surface. The light transmitted by this surface is confined within a bundle consisting of a chief ray with others round it not necessarily intersecting the chief ray. These rays are all of them normals to the wave front and are confined within a small volume round the chief ray. The wave front passes through the refracting surface, the shape is changed, the direction of the chief ray is changed and the formulae of the paper concern the imagery that is produced by the refracting surface.

It is presumed that the wave front is not spherical on either side of the re-

fracting surface which is in itself not spherical. The small part of any curved surface can be represented in the roughest form of approximation by the tangent plane and if the refraction be dealt with on these lines nothing further can be obtained than the line of emergence of the chief ray. The determination of conjugate points requires a closer degree of approximation to the shape of the wave front and to the shape of the refracting surface.

The next closer approximation to the wave front shape requires a knowledge of the two principal radii of curvature and of the principal planes. They are envisaged most easily by cutting off a slice from the surface parallel to the tangent plane. If the wave front is spherical the section is circular. If it is aspherical the section is a small ellipse unless the shape is saddle-backed when the section is a small hyperbola. The two principal radii of curvature are found by making sections of the surface through the normal and through the major axis for one value and the minor axis for the other. The centres of curvature are two quite distinct points along the chief ray and it is only rays in the principal planes that intersect the chief ray at all.

Now passage through the refracting surface not only changes the direction of the chief ray but it imposes new curvatures upon the wave front, and turns the principal planes into new directions. But any computation that has made this approximation to the shape of the wave fronts can be no more than a paraxial one. It will determine the two principal foci of the emergent wave front and will determine the directions in which the two focal lines must be drawn, but it will *not* determine anything at all about the shape of the caustic surface. We can only deal with the latter when we have a closer approximation to the shape of the surfaces, both the front and the refracting surface, than is dealt with by Mr Smith on p. 134. There, the formula $2x = Ay^2 + 2By^2 + Cz^2 + \ldots$ deals only with the two principal radii of curvature unless the coefficients of the higher powers on the right-hand side are included. The formulae are paraxial and cannot be used to determine aberrations.

But they can be used to determine the paraxial position of any skew ray since each normal round the indicatrix intersects each of the focal lines and the position of the line of shortest distance between the skew ray and the chief ray is easily determined. The difficulty that Mr Smith mentions on p. 135 to determine the extent of the astigmatism arises from the fact than an approximation to the shape of the wave front and to the shape of the refracting surface do not go far enough.

Mr Twyman: Investigations of general character in connection with the formation of images, and indeed concerning any subject, are not as a rule inspired by the immediate needs of the practitioner. Yet it is to such general investigations alone that we can look with hope for the discovery of any general principles of guidance.

Mr Hasselkuss has told us with what repugnance the computer brings himself to have anything to do with skew rays. Yet it is an infinitesimal proportion of the rays which are *not* skew. However much the computer may object to them it is the skew rays which form the image in an infinitely greater proportion than do rays which pass through the axis. General conclusions concerning the behaviour of

these rays can only be looked for from such enquiries as those of Mr Smith; such conclusions, for example, as might enable the computer to know with certainty when he must and when he need not trace skew rays.

Mr T. Smith (communicated): Prof. Sampson mentions the pitfalls presented by reflection. Although the path in this special case is quite direct, I agree that the unwary traveller is almost certain to have a fall on it. The difficulty has not been appreciated by most of our writers. The substitution $\mu = -1$ in refraction formulae is worthy of a chapter to itself in any critical treatise on optics.

Captain Baker has entered a plea for the computer. I should be sorry for the absence of numerical illustrations to be regarded as an indication of neglect of optical computers. I think, however, that the addition of one or two numerical examples would be of insufficient value. If new methods are to be adopted in such work, an exposition of much greater length than would be appropriate for inclusion in these *Transactions* would be desirable. There is ample ground to justify a book on such numerical work. There is, however, the further objection that it is not the immediate aim of this paper to recommend methods for numerical work. The determination of the algebraic forms appropriate for dealing with skew pencils seemed to me sufficient matter for a single paper. While I am in agreement with most that Captain Baker says, I think it is misleading to consider a single wave front or a single train of waves in the way he has done, because the theory deals with every wave front normal to the chief ray. The point is important; it is at the root of the difficulty concerning the bounds of astigmatism as considered in the paper. Captain Baker's final suggestion that a failure in approximation is involved is not correct, as I think he will realise when he has had an opportunity of studying the paper. The terms affecting simple astigmatism in the pencils considered are only of the second order and are thus taken fully into account.

No one wishes to trace skew rays as part of the routine of the computing office. To obviate this they must be studied in other ways. Until our knowledge is sufficiently extensive for us to be certain that we may neglect them in every system with impunity, we should aim at extracting the utmost information from every skew trace. Analysis on the lines given in the paper will enable information to be derived from a single skew ray that would otherwise only be obtainable by tracing a group of five skew rays, each computed with a much larger number of significant figures than is needed for the single skew ray.

We ought not to allow the great preponderance of axially symmetrical systems among manufactured optical instruments to let us lose sight of the importance of systems not having this symmetry. Asymmetrical systems are becoming of increasing importance in the physical laboratory. The human eye is often markedly unsymmetrical, and there is no other instrument so fully deserving of our attention and study. The time may come when the maker of normally symmetrical instruments will be expected to compensate to the utmost advantage particular samples for users whose vision is abnormal. All designers will then need to know more than is now required about unsymmetrical pencils.

CHARTS FOR SIMPLE TWO- AND THREE-LENS PROBLEMS

By T. SMITH, M.A., F.Inst.P.
(Optics Department, The National Physical Laboratory)

MS. received, 10*th December*, 1929. *Read and discussed*, 10*th April*, 1930.

ABSTRACT. A variety of charts can be drawn each of which furnishes complete first order information on systems constructed from two or three thin lenses.

Introduction

The first order relations in the theory of the symmetrical optical instrument are so simple that direct calculation is likely to be the general method of solving any problems that may arise. It will usually be found that graphical methods increase rather than diminish the labour involved. Possible exceptions in favour of graphical methods are found where a single chart gives the solution to a wide range of problems and exhibits at a glance the significance of many relations which are less immediately grasped from a set of equations. That a very wide field may be covered by a single chart is shown by the problem considered here, which deals with the construction of alternative forms of diagram for the solution of two-lens problems of the eye-piece type and three-lens problems of the photographic lens type.

The Two-lens Problem

Let A, B, C, D be the four Gaussian coefficients of the system, A being the power, B the magnification for an image situated at the reference point in the first medium, C the magnification for an image situated at the reference point in the final medium, and $-D$ the apparent separation of these reference points. The system is equivalent to two thin lenses placed at reference points a distance $-D$ apart, of powers $(C-1)/D$ and $(B-1)/D$ respectively. If the system consists of two thin lenses separated by a distance $-d_1$, the initial and final reference points being at distances $-d_0$ and $-d_2$ from the first and second lenses respectively (the distances being positive if measured away from the optical system), movement of the reference points to the positions of the actual lenses involves the replacement of

$$\begin{pmatrix} B & D \\ A & C \end{pmatrix}^{*} \text{ by } \begin{pmatrix} 1 & -d_0 \\ . & 1 \end{pmatrix} \begin{pmatrix} B & D \\ A & C \end{pmatrix} \begin{pmatrix} 1 & -d_2 \\ . & 1 \end{pmatrix}.$$

* For explanation of the notation and of the multiplication of matrices see *Trans. Opt. Soc.* **29** 1927–28 (167).

That is to say

$$1 + Ad_1 = (B - Ad_0)(C - Ad_2),$$

$$\frac{a_1}{A} = \frac{C - 1 - Ad_2}{Ad_1},$$

$$\frac{a_2}{A} = \frac{B - 1 - Ad_0}{Ad_1},$$

where a_1 and a_2 are the powers of the two thin lenses. If we write $-\delta$ for the separation of the original reference points

$$A\delta = A(d_0 + d_1 + d_2) = (B - 1 - Ad_0)(C - 1 - Ad_2) + B + C - 2.$$

The problem for solution will often be to determine a_1, a_2, d_0, d_1, d_2 from given values of A, B, C, D, δ. Let us write

$$Ad_0 = B - 1 + x,$$

$$Ad_1 = y,$$

$$Ad_2 = C - 1 + z,$$

$$A\delta = B + C - 2 + v,$$

$$\frac{a_1}{A} = u,$$

$$\frac{a_2}{A} = w,$$

where A merely determines the scale on which the diagram is to be interpreted. The additive quantities $B - 1$, $C - 1$, $B + C - 2$ are invariable in any given problem, and only affect the numerical value associated with a given curve of the chart. The equations then show that the solutions may be read from a chart exhibiting simultaneously the six families of curves, corresponding to constant values of u, v, w, x, y, z respectively, where

$$uy + z = 0,$$

$$v = x + y + z = xz,$$

$$wy + x = 0.$$

As the six variables are connected by four equations there are two degrees of freedom, and a plane diagram exhibits all possible relations between them. One of these variables will usually be determined directly from the imposed conditions, and there is then one degree of freedom left. The object of the chart is to show how much liberty this degree of freedom in fact allows, and how it may be most satisfactorily used.

The Three-lens Problem

The system A, B, C, D is equivalent to a single lens of power A with the reference points at distances $(1 - B)/A$ and $(1 - C)/A$ from the lens respectively. Let us suppose that the actual system consists of three lenses of powers a_1, a_2, a_3 respectively successively separated by distances $-d_1$ and $-d_2$, the reference points

being situated at the centres of the extreme lenses. By adding at the two ends lenses of powers $-a_1$ and $-a_3$ respectively the system degenerates effectively to the central lens alone, or

$$\begin{pmatrix} 1 & . \\ -a_1 & 1 \end{pmatrix} \begin{pmatrix} B & D \\ A & C \end{pmatrix} \begin{pmatrix} 1 & . \\ -a_3 & 1 \end{pmatrix} = \begin{pmatrix} 1 + a_2 d_1 & D \\ a_2 & 1 + a_2 d_2 \end{pmatrix},$$

or

$$1 + Da_2 = (C - Da_1)(B - Da_3),$$

$$\frac{d_1}{D} = \frac{B - 1 - Da_3}{Da_2},$$

$$\frac{d_2}{D} = \frac{C - 1 - Da_1}{Da_2}.$$

If α is written for the sum of the powers of the individual lenses

$$D\alpha = D(a_1 + a_2 + a_3) = (B - 1 - Da_3)(C - 1 - Da_1) + B + C - 2.$$

Obviously D determines the scale and the chart applicable to the previous problem may be used without alteration, the interpretation being

$$Da_1 = C - 1 + x,$$

$$Da_2 = y,$$

$$Da_3 = B - 1 + z,$$

$$D\alpha = B + C - 2 + v,$$

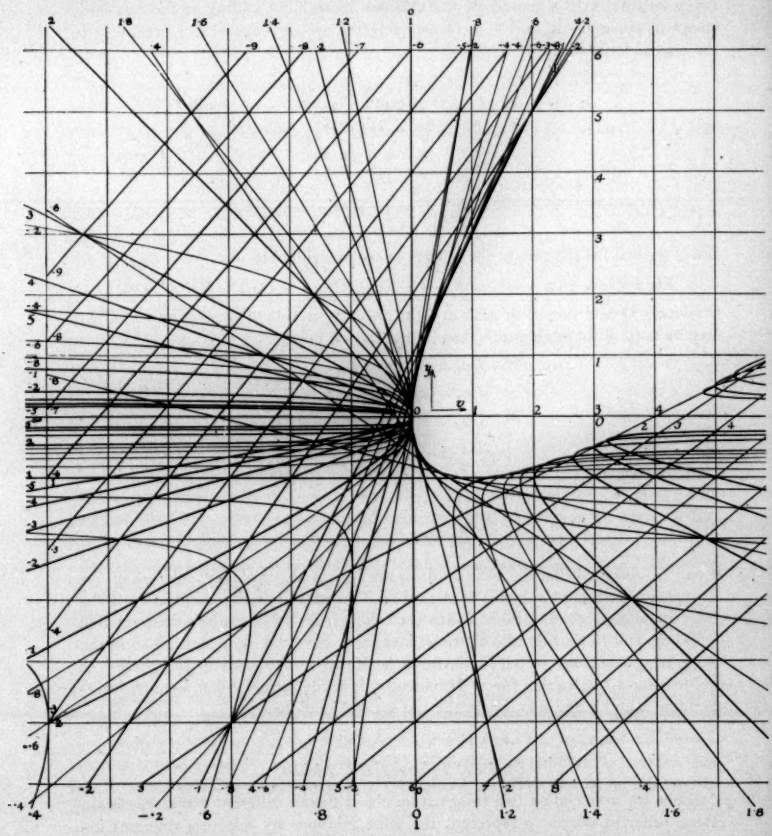
$$\frac{d_1}{D} = u,$$

$$\frac{d_2}{D} = w.$$

It will be observed that the two problems are in fact identical: any solution of one is, in theory at least, a solution of the other. The only limitation in practice to the transformability of one problem into the other arises from the requirement that d_1 and d_2, and frequently also d_0, must invariably have the same sign. Also in the two-lens problem we usually seek for a large value of δ, and in the three-lens problem for a small value for α. The identity is made formally complete by interchanging x and z and u and w in the last set of equations.

Choice of Independent Variables

Since we may choose two things from six in fifteen different ways, we obtain fifteen different charts to represent the same relations by selecting different independent variables. Twelve of these may be arranged in pairs, the two members of each pair only differing from one another in the labels attached to the curves. There are thus nine essentially different charts possible. They are far from being on an equality with one another in the ease with which they may be constructed. We differentiate between them by requiring as many curves as possible to be straight lines, and the remainder preferably of the second order. Discrimination is readily



The independent variables y and v are measured vertically and horizontally respectively, the scales being in the upper right-hand part of the figure. The oblique straight lines represent x and z, and the numerical values are inserted slightly within the figure. The hyperbolas represent u and w, and their values are outside the figure or at the extreme edge. The curves for fractional values of the variable are represented by fainter lines than those for integral values. When $u=0$ the hyperbola degenerates to $v=0$ with another branch at infinity, and when $u=1$ to $v=0$ and $y=-1$.

made, for example, by constructing the twenty equations connecting the variables three at a time.

The original four equations show that, if x and z are taken as independent variables, v and y, and hence also u and w, are represented by rectangular hyperbolas. Quite apart from the labour involved in plotting four series of such hyperbolas, such a chart would be very likely to be misread, for the asymptotes of all four families lie in parallel directions, and confusion would be difficult to avoid. Moreover the curves would in general intersect at small angles, and accurate reading would be difficult.

A better choice would be to adopt y and z as independent variables. Evidently u and x are then represented by straight lines, and v and w by hyperbolas. The straight lines u radiate from the origin, the straight lines x from $(-1, 1)$, the asymptotes of v are $z = 0$, $y + z = v$, those of w are $wy + 1 = 0$ and $wz - w + 1 = 0$. All the w curves pass through the radiant points of the u and x curves. This choice of variables, though unsymmetrical, is very convenient optically. For example, the overall length and the eye clearance are important factors in the design of telescopes.

A symmetrical system involving only two systems of conics is obtained by choosing u and w as independent variables. We readily deduce the equation

$$uwy + u + w - 1 = 0$$

from the original four. This shows that the y curves are rectangular hyperbolas, but x and z, i.e. uy and wy, are straight lines. The v curves are evidently given by

$$uvw = (u + w - 1)^2$$

and are also of the second order. They are closed if v lies between 0 and 4. These curves have been described by Mr T. Chaundy*.

In some respects the most favourable choice is v and y. Evidently x and z are then merely two members of the same system of straight lines, viz.

$$v(1 - x) + yx + x^2 = 0,$$

and u and w two members of the same family of conics, viz.

$$vyu + y^2u(u - 1) + v = 0,$$

which all pass through the origin. The typical straight line is most readily constructed by joining the points $(x, -1)$ and $(0, -x)$, and the hyperbola, which has the asymptotes $vu + (yu - 1)(u - 1) = 0$ and $yu + 1 = 0$, by drawing the curve through the points of intersection of x with $y = -x/u$. Through any point of this diagram (a certain area excepted) two of these straight lines and two of the hyperbolas pass. They represent x and z, and u and w, respectively. There is no difficulty in associating them correctly since $ux = wz$. This chart is shown on the opposite page.

It may be readily shown that any other choice of variables (except x, y, which is essentially the same as y, z) is less favourable than those considered above.

* "Two-lens combinations," *The Optician*, 31 Jan. 1919, p. 276.

Examples

(1) A telescope with an objective of 40 cm. focal length is to yield an angular magnification of − 10 with the image of distant objects 25 cm. in front of the eye, the distance from objective to Ramsden circle to be 50 cm.

The first step is to find the value of A, the power of the eye-piece. If we take the reference points at the objective and at the Ramsden circle we shall have $D = 0$, since these are conjugate points. Also, since angular magnification is inversely as linear magnification $B = -10$, $C = -0{\cdot}1$. From the distances -40 cm. and -25 cm. of the conjugate image planes we then find

$$A = \frac{-10}{-0{\cdot}4} + \frac{-0{\cdot}1}{-0{\cdot}25} \text{ dioptres}$$

$$= 25{\cdot}4 \text{ dioptres.}$$

Then
$$A\delta = 25{\cdot}4 \times (-0{\cdot}5) = -12{\cdot}7$$

$$= B + C - 2 + v = -10 - 0{\cdot}1 - 2 + v,$$

or
$$v = -0{\cdot}6.$$

Since A is positive we must obtain a solution with a negative value of y to secure a real separation between the two lenses of the eye-piece, and also $-d_2$ must be large enough to allow the eye to be placed correctly with respect to the eye lens. As an extreme value we may take 1 cm., giving

$$-0{\cdot}25 = -0{\cdot}1 - 1 + z,$$

or z may not exceed 0·85. It will probably be undesirable to adopt a value less than 0·35, corresponding to a separation between eye lens and Ramsden circle of 3 cm. By considering the range of the curve $v = 0{\cdot}6$ intercepted between $z = 0{\cdot}35$ and $z = 0{\cdot}85$ we see at once what latitude is available in the positions and powers of the two eye-piece lenses. The latitude will often be employed to enable standard parts to be utilised. The diagram also gives an indication of the precision necessary in working to the theoretical prescription for every variable involved.

(2) The design of photographic lenses of the Taylor type is to be considered with the extreme lenses distant 0·8 of the focal length from the corresponding principal foci. To ensure approximate observance of the Airy-Petzval condition the sum of the powers of the individual lenses is to be zero.

The conditions in terms of the quantities in the equations are

$$B = C = 0{\cdot}8, \qquad \alpha = 0,$$

hence
$$v = 0{\cdot}4.$$

From the identity between A, B, C, D, we have $AD = -0{\cdot}36$. It is a mere matter of convenience whether we adopt unity as the power of the lens, with $D = -0{\cdot}36$, or simplify the equations by putting $D = -1$, so that $A = 0{\cdot}36$, or make another choice. Obviously u and w must be positive, and therefore the sign of y must differ from those of x and z. No value of α which would make v negative is permissible.

DISCUSSION

Mr E. T. Hanson: When we are dealing with thin lenses close together the power of the combination is simply the sum of the powers of each component thin lens. The problem dealt with by Mr Smith can obviously include cases in which each "lens" is defined as any combination of thin lenses close together, so that such a combination may be denoted by the term "thin lens." Mr Smith has described a simple chart which shows the permissible limits between which the power, say, of each "thin lens" must lie in order to comply with the initial specification of the complete system. Having determined these limits, the subsequent modifications are obtained from other considerations, such as the approximate elimination of one or more aberrations or, as Mr Smith points out, the possibility of utilising standard parts. His chart is remarkable for its elegance and for the wide range which it covers.

Perhaps I may be pardoned for alluding to a paper [*Trans. Opt. Soc.* **25** (1923–24) 1] in which charts were used with some success in the discussion of a photographic lens, starting with "thin lenses" as just defined. In this case there are involved in each "thin lens" a number of variables which are capable of adjustment in order to insure, say, the correction of the first order aberrations. These variables are additional to those considered by Mr Smith.

There is little doubt that between the very special lens problem and the very general one there lies a wide region in which charts enable one to obtain a comprehensive survey which it is by no means easy to obtain in any other way.

Mr T. Smith (communicated): I am of the same opinion as Mr Hanson on the value of charts in exhibiting relations over a wide area. The chief limitations are the small number of independent variables we can admit and the restricted accuracy of graphical methods. As the drawing may occupy a considerable time it is important to make the field each chart covers as large as possible.

The problems considered here only concern first order relations, and are altogether more elementary than those under discussion in Mr Hanson's paper.

EXHIBIT OF A NEW SYSTEM OF TEST TYPE ILLUMINATION

By Messrs RAPHAEL'S Ltd.

Exhibited, 16th January, 1930.

A problem in connection with the design of ophthalmic equipment which does not appear, in many instances, to have been studied with very satisfactory results, consists in the provision of a uniform system of standard lighting for test types.

The difficulties of providing uniform front surface illumination in the case of test types of the drum form probably provide the only substantial argument in favour of the translucent panel type with internal lighting.

The writer has frequently heard the opinion expressed that the opaque panel type, having two or more internal lamps situated at or near the axis of the drum, constitutes an arrangement in which the illumination, when viewed externally, is as nearly perfect with respect to uniformity as is necessary. It does not fall within the scope of this brief survey of the matter to discuss the exact degree of uniformity which may, under various circumstances, be regarded as necessary. It may, however, be pointed out that the illumination in the case first considered is very far from being uniform. Not only may this be demonstrated by readings taken at various points on the surface by means of a photometer, but, from a brief theoretical consideration, it will be clear that a greater apparent intensity of light will occur over the region lying nearest to the line connecting the eye and the source of light. This effect may, of course, be remedied, to some extent, by the use of frosted lamps, and still further reduced by the provision of one or more translucent shells situated between the lamps and the inner surface of the test type.

The requirements of uniformity can, therefore, in this way and, further, by a multiplicity of lamps, be met to the required degree. The opaque test panels are, however, comparatively costly and fragile and are not readily interchangeable. One is indeed, therefore, induced to consider further the claims of the externally illuminated chart.

According to the recommendations of a Committee, appointed by the Council of British Ophthalmologists in July 1918, the source of illumination may consist of two lamps situated fifteen inches from the plane of the chart on the ends of arms arranged to position the lamps one-third of the chart's length up from its bottom edge and one-third down from its upper edge on the two sides. The arrangement provides a reasonably good result so far as uniformity of the illumination is concerned, but it is definitely not pleasing in appearance and, furthermore, the considerable projection of the lamp arms from a position on the wall at a height of less than about seven feet is objectionable.

This is the system generally employed for external illumination. With the object of reducing this distance of projection of the lamps, a multiplicity of lamps may be used, and, if evenly spaced along the length of the chart, four lamps will be found to give almost perfect uniformity in the vertical meridian, when placed at a distance of eighteen inches, and eight lamps when at a distance of twelve inches from the face of the chart.

This result may be achieved by the use of strip lights contained within a lamp tube of length approximately equal to the chart. With this arrangement of the light source, almost perfect illumination in the vertical meridian may be obtained. The proximity of the lamps to the surface does, however, introduce a certain unevenness of illumination in the horizontal meridian which is practically negligible when the lamps are situated at more than three feet from the surface. In decreasing this lies the chief problem of utilising light sources placed comparatively close to the surface, since the lamps must, in practice, be situated some distance outwards from the edge of the chart in order to provide an unobstructed view of the chart over a small range of angularity of view.

Simple calculations show, of course, that a curved form of chart will overcome the inequality of illumination in the horizontal meridian. A photometric examination of a surface curved in accordance with the theoretical investigations does not, however, show perfect uniformity, due partly to the fact that no account is taken of obliquity of vision and partly to the tendency of the surface, although of rough, unpolished material, to direct reflection. Any theoretical investigation of these last two modifying circumstances is difficult and hardly warranted in view of the facility with which a purely practical solution can be arrived at, with considerable accuracy, by which means it may be found that a shallow concave circular curvature of the front surface in the horizontal meridian will give a uniformity which, when viewed photometrically at the usual test distance, leaves nothing to be desired.

The use of a curved surface calls for any consequential "fore shortening" to be considered. Taking the outside letter of the widest (6/18) line, not a critical line, of a Snellen's type, the horizontal reduction of the letter is ·03 inch. On the 6/6 line—a critical line—the outside letter shows only ·013. These minute amounts, which even then only apply to the outside letters, may be ignored in the practical use of the test chart.

Notwithstanding the standardisation of intensity of illumination which may be expected to be established in the future, it is desirable for individual practitioners

to be able to vary this within wide limits for investigating the visual acuity as affected by such variation. A simple form of rheostat may, therefore, with advantage, be incorporated.

A calibration of the rheostat dial in units of illumination is not a suitable method of accurately determining the illumination used. Experiments which the writer made have suggested methods for overcoming the difficulties.

The matter of intensity of illumination as well as the colour of the light used is one calling for a standard to be set up by some authoritative body. My firm, pending the publication of such a standard, has decided upon an intensity and colour which, it is hoped, because of their closeness to conditions in nature, will be generally adopted and thus form the basis of future standards.

DISCUSSION

Dr J. W. T. Walsh: If I may speak as one interested in illumination problems, but without any knowledge of ophthalmic work, I should like to say that I have been very interested in Mr Raphael's exhibit. It has always seemed to me that the degree and evenness of the illumination of a test chart could not fail to have considerable influence on a subject's performance, and I have always therefore been surprised that no definite standards have been laid down in this respect. Mr Raphael's system would appear to provide a very satisfactory illumination from the point of view of uniformity, and the intensity provided would seem to be ample for the purpose, though, of course, it is considerably below what would be obtained anywhere near the window on a bright day. Until some standard is agreed upon, however, it seems useless to talk about adequacy or inadequacy. Would it not be possible for the Optical Society to take the matter up with either the Ophthalmic Society or with the Medical Research Council, who have a Committee to deal with problems connected with vision?

Mr O. P. Raphael (communicated): Dr Walsh's commendation is much valued.

The adoption of the standard of illumination which has been displayed was influenced by the following fact:

The patient would not require the intensity of normal daylight since he would usually have been tested and for some little time would have been in a considerably darkened room.

Further research in this matter would be welcomed so that a universally acceptable standard can be formulated, and in this connection your attention might be drawn to the Research Committee of the Institute of Ophthalmic Opticians, on which we know your Secretary, Mr A. Whitwell, sits in an advisory capacity.

EXHIBIT OF SCIENTYPE OCCLUDING GLASSES

By Messrs RAPHAEL'S Ltd.

Exhibited, 16th January, 1930.

An indispensable step in the treatment of squint is the improvement of vision in the amblyopic eye through occlusion of the fixing eye.

Prior to the introduction of Scientype Occluding Glasses, there were three methods of occlusion available:

(*a*) By means of opaque shields.

(*b*) By means of ground glass.

(*c*) By means of atropine.

All of these methods possessed disadvantages. Opaque shields, whilst extremely effective, are unsightly in wear and their use is often objected to on this account by parents. The appearance of ground glass, whilst an improvement upon that of opaque shields, is also objectionable. The instillation of atropine is a very fine method, but few optical or medical practitioners care to use it continuously over the long period generally necessary for improvement in vision to take place.

Scientype Occluding Glasses have one surface regularly patterned with depressions. This patterning is so arranged that the vision of the eye behind such a glass is reduced whilst the appearance of the eye from the front is unaltered. The patterning can be controlled to give any degree of occlusion within widely separated limits. It has been found that two grades are sufficient.

Grade A. Reducing V from 6/6 to 6/60.

Grade B. Reducing V from 6/6 to 1/60.

The degree of occlusion in any particular case is governed by the degree of V. A. possessed by the deviating eye. In order to force the deviating eye to make effort to see, it is necessary merely that the V of the fixing eye shall be reduced below that of the deviating eye.

The occluders have recently been improved in two respects. First, they are now made in 9D base meniscus form to increase the occlusion field. Secondly, they are made by a new process which ensures that the occlusion effect is constant over the whole glass area. Unless the effect be thus constant, it is found that child wearers contrive to look through one particular point where better than average vision may be obtained.

This form of occluding glass also possesses the advantage of being safer in use, since it permits of large objects—such as traffic—being seen, and because it is free from any tendency to dazzle. Further, the occlusion effect is controllable, since the glass may be worn, or not, according to the need of the moment.

REVIEW

The Kinematical Design of Couplings in Instrument Mechanisms. By Professor A. F. C. POLLARD, A.R.C.S., A.M.I.E.E., Pp. 64. (London: Adam Hilger, Ltd.) Price 4*s*. 6*d*. net.

Under this rather formidable title Prof. Pollard, the professor of instrument design at the Imperial College of Science and Technology, has written a monograph of sixty-four pages on the geometrical design of scientific instruments.

Prof. Pollard is well known as an exponent of geometrical design as applied to instrument mechanisms, and there is little doubt that as a result of his Cantor Lectures in 1922 the professional makers of scientific instruments are using kinematical design to a much greater extent than formerly. That there was urgent need for the introduction of carefully considered mechanisms in many instruments—especially those used micrometrically—was felt by many users. As a whole the instrument manufacturer is a conservative individual, and it has taken a long time to persuade him that kinematic design may improve the efficiency of his product, and will undoubtedly reduce the cost of its manufacture.

The first twenty-three pages of the book are devoted to a theoretical consideration of the type and magnitude of the errors found in measuring instruments. The author divides these instruments into three groups: indicating, value controlling and counting mechanisms. On reflection it is interesting to note how the instruments in daily use fall into one or other of these three categories. In the book consideration is given almost exclusively to the first two types.

Prof. Pollard devotes some space to defining *variance* and in separating it from *accuracy* and *sensitivity*. In practice it is often difficult to do this. The word *passiveness* is used for defining the smallest change in the measured quantity which "will produce any change whatever in the indication of the instrument." This is undoubtedly a useful piece of information in the case of some instruments; for example, the amount of current required to deflect the coil in a moving coil galvanometer when this is being used in a "null method" circuit.

In the introductory pages Prof. Pollard defines what is understood by "six degrees of freedom," and arranges various familiar mechanisms by the degrees of freedom they possess.

The remainder of the book is devoted to examples of kinematic design as applied to scientific measuring instruments. Some of the designs are beautiful in their simplicity, the best criterion of good design. The efficiency and simplicity of the design of the mounting for the mirrors and parallel plate of an interferometer is an example in question.

The use of balls in enabling two members of an instrument to move relatively to each other with the minimum of friction is just coming into general use, although there were one or two outstanding examples designed nearly thirty years ago. The ball races in the measuring machine designed by Prof. Dalby and shown in Fig. 22, are a good example of modern design. Ball bearings are now being largely used in heavy astronomical instruments. It is a good example of the re-actance between science and industry. Science has helped the manufacturer of bearings to make balls of a high degree of accuracy: the scientific man now takes these balls to make instruments capable of still higher accuracy.

We wish Prof. Pollard could have found space to illustrate some other well-known geometric mechanisms. After reading the book we have to confess to a feeling of disappointment. A book is wanted which will enable the designer to realise the importance of the geometrical principles of design both with a view to increasing accuracy and also reducing the labour required to make an instrument for a given purpose. We had hoped that in this monograph Prof. Pollard would have done this. The introduction assumes that the reader has the mathematical type of mind, and therefore plunges him into definitions of the degrees of freedom etc. Two or three simple diagrams would have brought home the fundamental ideas underlying the meaning of kinematical design and would, we believe, enable the student to think round the subject more clearly. The average draughtsman is a modest person, and we are afraid will fight shy of what is a useful handbook simply because the opening pages are rather too "high-brow." Those who are interested in design will read the monograph with pleasure, even if they are inclined to question some of the statements in the introductory pages, and will wish that as a mathematical treatment of the subject it had been more complete.

The community owes Messrs Adam Hilger, Ltd., a debt for publishing the book so well and at such a low price.

R.S.W.